

**TITLE: ASSAYS, METHODS AND APPARATUS FOR ASSESSING RNA
DISRUPTION**

FIELD

- 5 **[0001]** The various embodiments described herein generally relate to assays, methods and apparatus for assessing RNA disruption optionally in response to a cytotoxic treatment optionally chemotherapy and/or radiation therapy.

BACKGROUND

- 10 **[0002]** Ribonucleic acids (RNA) are biopolymers which encode genetic information and play various roles in a cell, including encoding proteins. RNA preparations are employed in the investigation of gene expression, for example, by microarray experiments, RT-PCR and many other methods. The results of experiments employing RNA preparations and the significance of results obtained by such
- 15 experiments, is largely dependent upon the integrity of the RNA employed. Different methods are available to measure RNA integrity of a sample. Typically the methods compare heat degraded or RNase degraded samples to samples with intact RNA with regard to the capacity of the RNA at a particular degradation level to be sufficiently intact to permit PCR amplification of specific mRNAs such as "housekeeping" genes.
- 20 **[0003]** For example, Schroeder, A., O. Mueller, et al. (2006) describes a method that automatically selects features from signal measurements and constructs regression models based on a Bayesian learning technique. Feature spaces of different dimensionality are compared in the Bayesian framework, which allows selecting a final feature combination corresponding to models with high posterior probability. The
- 25 approach was applied to a large collection of electrophoretic RNA measurements recorded with an Agilent 2100 bioanalyzer to develop an algorithm that describes RNA integrity. The resulting algorithm is a user-independent, automated and reliable procedure for standardization of RNA quality control that allows the calculation of an RNA integrity number (RIN) under certain conditions and/or for certain samples.

- 2 -

[0004] A method of using tumour RNA integrity to measure response to chemotherapy in cancer patients is disclosed in PCT/CA2008/001561 filed September 5, 2008.

5 **SUMMARY OF VARIOUS EMBODIMENTS**

[0005] In one aspect, in at least one embodiment described herein, there is provided a method for performing an RNA Disruption Assay (RDA) for cellular RNA optionally in response to a cytotoxic treatment optionally a treatment for a proliferative disorder optionally a cancer treatment, optionally chemotherapy, adjuvant or radiation
10 treatment. The method comprises obtaining at least one electropherogram dataset corresponding to a unique biological sample comprising cellular RNA optionally at a time point before, during or after the treatment; determining values for features from the at least one electropherogram dataset by using two identifying ranges to accommodate possible shifting of 18S and 28S peaks, detecting the 18S and 28S peaks, and
15 calculating the features derived at least in part based on the located 18S and 28S peaks; and optionally determining an RDA score based on a combination of the values of the features.

[0006] In one aspect, in at least one embodiment described herein, there is provided a method for performing an RNA Disruption Assay (RDA) for cellular RNA
20 response to a cytotoxic treatment optionally chemotherapy treatment and/or radiation treatment. The method comprises obtaining at least one electropherogram dataset corresponding to a unique biological sample comprising cellular RNA at a time point before, during or after the treatment; determining values for features from at least two shifted regions of the at least one electropherogram dataset, the shifting being due to the
25 treatment; and optionally determining an RDA score based on a combination of the values of the features.

[0007] In another aspect, in at least one embodiment described herein, there is provided an assay comprising: quantifying the amount of RNA disruption in a biological sample comprising cellular RNA at a time point before, during or after cytotoxic
30 treatment, optionally a cancer treatment, optionally chemotherapeutic, adjuvant and/or

- 3 -

radiation treatment, using an RNA disruption assay (RDA) performed according to the method for performing RDA described herein; and determining if the amount of RNA disruption of the cellular RNA is increased or not increased.

[0008] In another aspect, in at least one embodiment described herein, there is provided an in vitro assay for determining sensitivity to a cytotoxic treatment optionally a cancer treatment e.g. chemosensitivity of cells to a chemotherapeutic treatment and/or optionally radiosensitivity of cells to a radiation treatment. The assay comprises quantifying the amount of RNA disruption in a biological sample comprising cell RNA at a time point before, during or after treatment using an RNA Disruption Assay (RDA) performed according to the method for performing RDA described herein; and comparing the quantified amount of RNA disruption with a threshold or reference value, and if the amount of RNA disruption is increased relative to the reference value, identifying the cell as sensitive (e.g. chemosensitive and/or radiosensitive) to the treatment and if not increased identifying the cells as resistant to the treatment.

[0009] In another aspect, in at least one embodiment described herein, there is provided an assay for determining if a subject is responding to a cytotoxic therapy, optionally a cancer therapy, optionally a chemotherapeutic drug and/or radiation treatment comprising: assaying a biological sample obtained from the subject before, during and/or after the subject has received the treatment e.g. chemotherapeutic and/or radiation treatment for the quantity of RNA disruption using an RNA Disruption Assay (RDA) performed according to the method for performing RDA described herein, wherein the subject is identified as responding to the treatment if the quantity of RNA disruption is increased relative to the threshold or reference value.

[0010] In another aspect, in at least one embodiment described herein, there is provided a method of treating a subject in need of cancer treatment comprising administering a cancer treatment, optionally a chemotherapeutic drug, adjuvant and/or radiation; determining if the subject is responding to the treatment according to the RDA method described herein; and continuing administration of the treatment if the subject is responsive; and/or altering and/or discontinuing administration of the treatment if the subject is not responsive.

- 4 -

[0011] In another aspect, in at least one embodiment described herein, there is provided a non-transitory computer readable medium comprising a plurality of instructions executable on a microprocessor of an apparatus for adapting the apparatus to implement a method for performing an RNA Disruption Assay (RDA) for cellular RNA optionally for a proliferative disorder optionally a cancer treatment, optionally chemotherapy, adjuvant and/or radiation treatment. The method comprises accessing at least one electropherogram dataset corresponding to a unique biological sample comprising cellular RNA at a time point before, during or after the treatment; determining values for features from the at least one electropherogram dataset by using two identifying regions to accommodate possible shifting of 18S and 28S peaks, detecting the 18S and 28S peaks, and calculating the features derived in part based on the located 18S and 28S peaks; and optionally determining an RDA score based on a combination of the values of the features.

[0012] In another aspect, in at least one embodiment described herein, there is provided an apparatus for performing an RNA Disruption Assay (RDA) for cellular RNA optionally a treatment for a proliferative disorder optionally a cancer treatment, optionally chemotherapy, adjuvant or radiation treatment. The apparatus comprises a data input to receive at least one electropherogram dataset corresponding to a unique biological sample comprising cellular RNA at a time point before, during or after the treatment; a processing unit coupled to the data input, the processing unit being configured to determine values for features from the at least one electropherogram dataset by using two identifying ranges to accommodate possible shifting of 18S and 28S peaks, detecting the 18S and 28S peaks, and calculating the features derived at least in part based on the located 18S and 28S peaks; and a data output coupled to the processing unit to convey an indication of the RDA score.

[0013] In another aspect, in at least one embodiment described herein, there is provided an apparatus for performing an RNA Disruption Assay (RDA) for cellular RNA optionally a treatment of a proliferative disorder optionally a cancer treatment, optionally chemotherapy, adjuvant or radiation treatment. The apparatus comprises means for obtaining at least one electropherogram dataset corresponding to a unique biological sample comprising cellular RNA at a time point before, during or after the treatment;

- 5 -

means for determining values for features from the at least one electropherogram dataset by using two identifying ranges to accommodate possible shifting of 18S and 28S peaks, detecting the 18S and 28S peaks, and calculating the features derived at least in part based on the located 18S and 28S peaks; and means for determining an
5 RDA score based on a combination of the values of the features.

[0014] In another aspect, in at least one embodiment described herein, there is provided a method for performing an RNA Disruption Assay (RDA) for cellular RNA. The method comprises obtaining at least one electropherogram dataset corresponding to a unique biological sample comprising the cellular RNA at a time point; defining an 18S
10 peak and a 28S peak from the at least one electropherogram dataset; determining at least one parameter value for both the 18S peak and the 28S peak; redefining at least one of the 18S peak and the 28S peak when required according to one or more rules applied to the at least one parameter value; determining an 18S peak area and a 28S peak area; and determining an RDA score based on at least one of the 18S peak area
15 and the 28S peak area.

[0015] In another aspect, in at least one embodiment described herein, there is provided a method for performing an RNA Disruption Assay (RDA) for cellular RNA. The method comprises obtaining at least one electropherogram dataset corresponding to a unique biological sample comprising the cellular RNA at a time point; defining an 18S
20 peak and a 28S peak from the at least one electropherogram dataset; determining at least one parameter value for both the 18S peak and the 28S peak; and redefining at least one of the 18S peak and the 28S peak when required according to one or more rules applied to the at least one parameter value.

[0016] In another aspect, in at least one embodiment described herein, there is provided a non-transitory computer readable medium comprising a plurality of instructions executable on a microprocessor of an apparatus for adapting the apparatus to implement a method for performing an RNA Disruption Assay (RDA) for cellular RNA. The method comprises: obtaining at least one electropherogram dataset corresponding to a unique biological sample comprising the cellular RNA at a time point; defining an
25 18S peak and a 28S peak from the at least one electropherogram dataset; determining at least one parameter value for both the 18S peak and the 28S peak; and redefining at
30

- 6 -

least one of the 18S peak and the 28S peak when required according to one or more rules applied to the at least one parameter value.

[0017] In another aspect, in at least one embodiment described herein, there is provided an apparatus for performing an RNA Disruption Assay (RDA) for cellular RNA.

- 5 The apparatus comprises a data input to receive at least one electropherogram dataset corresponding to a unique biological sample comprising cellular RNA at a time point; a processing unit coupled to the data input, the processing unit being configured to define an 18S peak and a 28S peak from the at least one electropherogram dataset; determine at least one parameter value for both the 18S peak and the 28S peak; and redefine at
- 10 least one of the 18S peak and the 28S peak when required according to one or more rules applied to the at least one parameter value; and a data output coupled to the processing unit to convey an indication of the RDA.

- [0018]** Other features and advantages of the present disclosure will become apparent from the following detailed description. It should be understood, however, that
- 15 the detailed description and the specific examples while indicating preferred embodiments of the disclosure are given by way of illustration only, since various changes and modifications within the spirit and scope of the disclosure will become apparent to those skilled in the art from this detailed description.

20 **BRIEF DESCRIPTION OF THE DRAWINGS**

[0019] For a better understanding of the various embodiments described herein, and to show more clearly how these various embodiments may be carried into effect, reference will be made, by way of example, to the accompanying drawings which show at least one example embodiment, and in which:

- 25 **FIG. 1** is an illustration of an electropherogram trace of RNA isolated from a punch biopsy of a MCF-7 tumour xenograft that has been kept at room temperature in saline for 24 hours where autolytic degradation takes place;

FIG. 2 is an illustration of an electropherogram trace of RNA isolated from human A2780 ovarian cancer cells that have been treated for 24 hours at 37°C with the

- 7 -

chemotherapeutic agent docetaxel (10 μ M in medium diluted 1 in 2 with phosphate buffered saline (50% media);

FIG. 3 is an illustration of an electropherogram trace of RNA isolated from A2780 cells treated for 24 hours at 37°C with the chemotherapeutic epirubicin (20 μ M in
5 50% media);

FIG. 4 is a block diagram of an example embodiment of an apparatus that can be used to assess RNA disruption;

FIG. 5A is a flowchart of an example embodiment of a method that can be employed to assess RNA disruption;

10 **FIG. 5B** is a flowchart of an example embodiment of a method that can be employed to determine features related to 18S and 28S peaks in the 18S and 28S shifted regions respectively;

FIG. 5C is a graph of an example electropherogram that defines the 18S peak, 28S peak, intermediate banding region, low banding region and the marker region;

15 **FIG. 6A** is a log-log graph of the feature of $\log(\text{Maximum Intermediate}/(28\text{S}+18\text{S}))$ versus $\log$ RNA concentration;

FIG. 6B is a semi-log graph of the feature of $\text{Minimum Intermediate}/(28\text{S}+18\text{S})$ versus $\log$ RNA concentration for a sample study;

FIG. 6C is a semi-log graph of the feature of $\text{Maximum Intermediate/Total Area}$ versus $\log$ RNA concentration for a sample study;
20

FIG. 6D is a semi-log graph of the feature of $\text{Minimum Value Low Banding/Total Area}$ versus $\log$ RNA concentration for a sample study;

FIG. 6E is a graph of the feature $\text{Low to Total Area ratio}$ versus $\text{Maximum Intermediate to Total Area ratio}$ for a sample study;

25 **FIG. 6F** is a semi-log graph of the feature of $\text{Maximum 28S/Total Area}$ versus $\log$ RNA concentration for a sample study;

FIG. 6G is a semi-log graph of the feature of $\text{Minimum 28S/Total Area}$ versus $\log$ RNA concentration for a sample study;

FIG. 6H is a semi-log graph of the feature of Maximum 18S/Total Area versus log RNA concentration for a sample study;

FIG. 6I is a semi-log graph of the feature of Minimum 18S/Total Area versus log RNA concentration for a sample study;

5 **FIG. 6J** is a semi-log graph of the feature of Maximum Intermediate/28S versus log RNA concentration for a sample study;

FIG. 6K is a semi-log graph of the feature of Minimum Intermediate/28S versus log RNA concentration for a sample study;

FIG. 7A is a graph of patient distribution in RDA zones;

10 **FIG. 7B** is an ROC curve corresponding to **FIG. 6A**;

FIG. 8A is an image of a gel of electrophoretically separated xenograft tumour RNA samples;

FIG. 8B is an illustration of an electropherogram trace of lane 1 of the gel of **FIG. 8A**;

15 **FIG. 8C** is an illustration of an electropherogram trace corresponding to lane 2 of the gel of **FIG. 8A**;

FIG. 8D is an illustration of an electropherogram trace corresponding to lane 3 of the gel of **FIG. 8A**;

20 **FIG. 8E** is an illustration of an electropherogram trace corresponding to lane 4 of the gel of **FIG. 8A**;

FIG. 8F is an illustration of an electropherogram trace corresponding to lane 5 of the gel of **FIG. 8A**;

FIG. 8G is an illustration of an electropherogram trace corresponding to lane 6 of the gel of **FIG. 8A**;

25 **FIG. 8H** is an illustration of an electropherogram trace corresponding to lane 7 of the gel of **FIG. 8A**;

FIG. 8I is an illustration of an electropherogram trace corresponding to lane 8 of the gel of **FIG. 8A**;

FIG. 8J is an illustration of an electropherogram trace corresponding to lane 9 of the gel of **FIG. 8A**;

FIG. 8K is an illustration of an electropherogram trace corresponding to lane 10 of the gel of **FIG. 8A**;

5 **FIG. 8L** is an illustration of an electropherogram trace corresponding to lane 11 of the gel of **FIG. 8A**; and

FIG. 8M is an illustration of an electropherogram trace corresponding to lane 12 of the gel of **FIG. 8A**;

10 **FIG. 9A** is a graph demonstrating the number of A2780 ovarian cancer cells in culture after treatment with different concentrations of docetaxel for different time periods;

FIG. 9B is a graph of the amount of RNA per cell for A2780 ovarian cancer cells treated with different concentrations of docetaxel for different time periods;

15 **FIG. 9C** is a graph plotting cell counts for cells treated with 0.2 micromolar docetaxel at different time points;

FIG. 9D is an image of a gel of electrophoretically separated A2780 ovarian cancer cell RNA samples treated with different concentrations of docetaxel for 24 hours;

20 **FIG. 9E** is an illustration of an electropherogram trace corresponding to 0.005 micromolar and 0.2 micromolar lanes of **FIG. 9D**;

FIG. 9F is a graph of the intermediate area/(28S+18S) calculated from the trace of **FIG 9E**;

FIG. 9G is a graph of the Low Area/(28S+18S) calculated from the trace of **FIG 9E**;

25 **FIG. 9H** is an image of a gel of electrophoretically separated A2780 ovarian cancer cell RNA samples of cells treated with different concentrations of docetaxel for different time periods.

FIG. 9I is a series of images of gels of electrophoretically separated A2780 sensitive and resistant cell RNA samples;

FIG. 9J is an image of a gel of electrophoretically separated A2780 ovarian cancer cell RNA samples treated with different concentrations of docetaxel for
5 different time periods in the presence or absence of caspase inhibitor;

FIG. 10A is a graph of the intermediate area/(28S+18S) calculated for A2780 ovarian cancer cells treated with radiation for different time periods;

FIG. 10B is an image of a gel of electrophoretically separated A2780 ovarian cancer cell RNA samples of cells treated with radiation for different doses and
10 time periods;

FIG. 11A is a graph of the intermediate area/(28S+18S) versus concentration calculated for patients treated with FEC, radiation and docetaxel;

FIG. 11B is a graph of the low Area/(28S+18S) for patients treated with FEC, radiation and docetaxel;

FIG. 12A is a flowchart of another example embodiment of a method that can be employed to determine features related to the 18S peak and the 28S peak in the
15 18S and 28S shifted regions respectively;

FIG. 12B is a graph of an example electropherogram in which the 18S peak, the 28S peak, the intermediate banding region, the low C banding region, the low
20 B banding region, the low A banding region and the marker region are defined;

FIG. 12C is a graph showing how a peak is determined using the method of **FIG. 12A**;

FIG. 13 is a log-log graph of the feature of $\log(\text{Intermediate}/(28\text{S}+18\text{S}))$ versus $\log$ RNA concentration for a sample study using the modified peak identification
25 method;

FIG. 14A is a graph of patient distribution in RDA zones using the modified peak identification method;

FIG. 14B is an ROC curve corresponding to **FIG. 14A**;

FIG. 14C is a log-log graph of the feature of $\log (\text{low } C/(28S + 18S))$ versus $\log$ RNA concentration for a sample study using the modified peak identification method;

FIG. 14D is a log-log graph of the feature of $\log (\text{Intermediate} + \text{low } C)/(28S + 18S)$ versus $\log$ RNA concentration for a sample study using the modified peak identification method;

FIG. 15A is a graph of the feature of (28S peak width + 18S peak width) versus (28S+18S) for a sample study using the modified peak identification method;

FIG. 15B is the graph of **FIG. 15A** further comprising Linear Discriminant Analysis (LDA) determined partition lines;

FIG. 15C is the graph of **FIG. 15A** further comprising another example pair of LDA determined partition lines;

FIG. 15D is the graph of **FIG. 15A** further comprising Quadratic Discriminant Analysis (QDA) determined partition lines;

FIG. 16A is a graph of the feature $\text{low } C/(28S+18S)$ versus docetaxel dose for an in vitro study using the modified peak identification method;

FIG. 16B is a graph of the feature $(\text{Intermediate} + \text{low } C)/(28S+18S)$ versus docetaxel dose for an in vitro study using the modified peak identification method;

FIG 17A is a graph of the feature $\text{Maximum Intermediate}/(28S+18S)$ versus RNA concentration for a sample study for patients undergoing neo-adjuvant therapy with Trastuzumab;

FIG. 17B is a graph of the feature $\text{Maximum Intermediate}/(28S+18S)$ versus RNA concentration for a sample study for patients undergoing neo-adjuvant therapy with Trastuzumab;

FIG. 17C is a graph of the feature $\text{Maximum Intermediate}/(28S+18S)$ versus RNA concentration for a sample study for patients undergoing neo-adjuvant therapy with Zometa;

- 12 -

FIG. 17D is a graph of the feature Maximum Intermediate/(28S+18S) versus RNA concentration for a sample study for patients undergoing neo-adjuvant therapy with Zometa;

FIGS. 18A–18F are a series of illustrations of electropherogram traces
5 with the calculated RIN and RDI values for each;

FIG. 19 is an image of a gel of electrophoretically separated RNA samples and corresponding RIN and RDI values for each sample;

FIG. 20A is a graph of the feature Intermediate Area / (28S + 18S Areas) for colon adenocarcinoma cells incubated with Sunitinib;

10 **FIG. 20B** is a graph of the feature Intermediate Area / (28S + 18S Areas) for colon adenocarcinoma cells incubated with Mafosfamide; and

FIG. 20C is a graph of the feature Intermediate Area / (28S + 18S Areas) for colon adenocarcinoma cells incubated with Epirubicin.

15 **DETAILED DESCRIPTION OF THE EMBODIMENTS**

[0020] Various assays, methods or apparatuses will be described below to provide an example of an embodiment of each claimed invention. No embodiment described below limits any claimed invention and any claimed invention may cover processes or apparatuses that differ from those described below. The claimed
20 inventions are not limited to apparatuses or processes having all of the features of any one apparatus or process described below or to features common to multiple or all of the apparatuses or processes described below. It is possible that an apparatus or process described below is not an embodiment of any claimed invention. Any invention disclosed in an apparatus or process described below that is not claimed in this
25 document may be the subject matter of another protective instrument, for example, a continuing patent application, and the applicants, inventors or owners do not intend to abandon, disclaim or dedicate to the public any such invention by its disclosure in this document.

[0021] Furthermore, it will be appreciated that for simplicity and clarity of illustration, where considered appropriate, reference numerals may be repeated among the figures to indicate corresponding or analogous elements. In addition, numerous specific details are set forth in order to provide a thorough understanding of the
5 embodiments described herein. However, it will be understood by those of ordinary skill in the art that the embodiments described herein may be practiced without these specific details. In other instances, well-known methods, procedures and components have not been described in detail so as not to obscure the embodiments described herein. Also, the description is not to be considered as limiting the scope of the embodiments
10 described herein.

I. Definitions

[0022] The term “coupled” used herein indicates that two elements can be directly coupled to one another or coupled to one another through one or more
15 intermediate elements.

[0023] The term “autolytic RNA degradation” as used herein refers to RNA degradation taking place during autolytic cell destruction. Autolysis is initiated, for example, by the cells' lysosomes releasing digestive enzymes into the cytoplasm due to the cessation of active processes in the cell, and not due to an active physiologic or
20 pathophysiologic process. Autolytic RNA degradation in a sample can be induced by removal of cells from physiologic environment for extended period of time (e.g. incubation in saline) and/or nonspecific cessation of physiologic processes e.g. heat treatment.

[0024] The term “RNA disruption” as used herein refers to discretely fragmented
25 and/or degraded RNA that is signal induced in response to a cytotoxic treatment such as a drug treatment, e.g. chemotherapy, radiation treatment, and/or cytotoxic antibody treatment (e.g. Trastuzumab). Cytotoxic signal induced RNA disruption can include RNA degradation that has some features that resemble autolytic degradation, particularly for example during later stages.

[0025] The term "subject" as used herein refers to any member of the animal kingdom, preferably a human being including, for example, a subject that has or is suspected of having a proliferative disorder such as cancer.

[0026] The term "control" as used herein refers to a comparator such as intact RNA, a pretreatment sample(s) from a subject or group of individuals with a known outcome and/or a value corresponding to or derived from such samples. With respect to methods for assessing treatment outcome, the control can be a reference value such as a baseline amount or a median or average pretreatment amount of RNA degradation for a particular cell or cancerous tumour. The control can be a sample from a subject or subjects who are known as responding or not responding to a cytotoxic treatment such as a cancer treatment e.g. chemotherapeutic and/or radiation treatment or a value determinable from such samples, such as one or more threshold values above which or below which (depending on how the threshold or measurement is defined) identifies the likelihood that a sample has disrupted RNA, or that a subject or disease is responsive to the cytotoxic treatment e.g. that the tumour is responsive to a chemotherapeutic treatment. The control can also be more than one control, such as a series of threshold values that define clinical zones for a known outcome group of subjects, wherein each zone is associated with a likely response. Untreated cells (e.g. cells pretreatment) from many cells types or, for example, from pretreated tumour samples can have a high RNA integrity. Accordingly, the control can, for example, be a pretreatment value derived from one or more pretreatment samples. As described below, subjects in the MA22 study with mid-treatment tumour values above a threshold (e.g. which exhibited increased degradation) correlated with a pathologic complete response within the breast and axilla post-treatment.

[0027] The term "internal standard" as used herein means an RNA sample that is used as a normalizer for a particular assay. For example, when the assay comprises using an RNA chip, the internal standard can be the sample that is determined as having the smallest value for the measure: $(\text{intermediate area} + \text{low B area} + \text{low C area}) / (\text{28S area} + \text{18S area})$. The internal standard can be used to identify which samples are adjusted. The internal standard can be a subject sample and/or a control sample.

[0028] The term "baseline amount" as used herein refers to an amount of RNA degradation (e.g. RNA disruption and/or autocatalytic RNA degradation) in a sample such as a pretreatment sample that is used for comparison to a test sample (e.g. comparison to a cell population and/or tumour) taken at a later time point, for example during or after treatment e.g. during or after a treatment regimen comprising chemotherapy, cytotoxic antibody and/or radiation treatment. For example, in methods related to monitoring response to treatment, "base-line amount" can refer to a level of RNA degradation in a sample taken prior to a subsequent sample, e.g. a base-line sample is taken before treatment, the comparison to which provides an indication of response to treatment.

[0029] The term "amount" as used herein with respect to RNA degradation (e.g. disruption) refers to an amount (e.g. relative amount or absolute amount) of RNA degradation/disruption that is detectable or measurable in RNA isolated from a sample, for example, using an RDA assay described herein. For example, the amount can be expressed using an absolute value (e.g. an RDI value) or a relative amount such as 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.2, 2.4, 2.6, 2.8, 3.0, 3.2, 3.4, 3.6, 3.8, 4.0, 4.2, 4.4, 4.6, 4.8, 5.0, 10, 15, 20, 25, 30, 40, 60, 80 and/or 100 times a control amount, where for example, the control amount is, for example, a pre-treatment amount or, for example, a reference value corresponding to the average or median level in untreated or pretreatment samples. The amount can be compared to a threshold value which identifies subjects with a likelihood of responding to the treatment.

[0030] The term "threshold value" with respect to discriminating subject outcome also referred to as "cut-off" or "cut-off value" as used herein refers to a value which is derived from a population of samples with known outcomes and identifies a clinical zone boundary (e.g. RDA zone) for a selected specificity and/or sensitivity. Multiple threshold values defining multiple clinical zone boundaries can be employed. The zone boundaries define clinical zones and subjects having an RDA score (e.g. an RDI) falling within a RDA zone are identified as having an increased likelihood of responding or not responding to a particular cytotoxic treatment e.g. chemotherapeutic and/or radiation treatment, defined by the zone. For example, the threshold value can be the highest value associated with a panel of known outcome patients (e.g. 100% specificity) or a

median level or other selected level, for example, as calculated according to one of the Examples described herein. A subject with an RDI score above the threshold value or falling within a zone specified as being below one threshold value or between two threshold values or above a second threshold value is predicted for example to be likely
5 responding to the treatment, indeterminate or likely not responding to the treatment. The threshold value can be derived, for example, from a database of a plurality of samples.

[0031] The terms "18S" and "28S" generally refer to the 18S peak area and the 28S peak area, respectively, unless these terms are modified by other words such as peak or width or band (which can refer to the 18S and 28S bands separated on a gel).
10 Accordingly, the terms "18S peak" and "28S peak" represent the heights of the 18S and 28S peaks respectively. Furthermore, the terms 18S width and 28S width represent the widths of the 18S and 28S peaks respectively. Once a peak is located, the peak width and the peak height are determined according to certain methods described herein. For example, the peak width may be defined by the subtraction of the starting time of the
15 peak from the ending time of the peak. Another example is that the peak height may be defined as the absolute value of the highest point in a peak region or band.

[0032] The term "specificity" as used herein refers to the percentage of subjects that are responding to a treatment that are identified as not responding to the treatment based on a RNA disruption score that is, for example, at or below a control level and/or a
20 cut-off level.

[0033] The term "sensitivity" as used herein refers to the percentage of subjects that are responsive that are identified as responding to the treatment based on a RNA disruption assay score, that is, for example, above a control amount and/or threshold value.

25 **[0034]** The term "sample" as used herein refers to 1) any biological fluid, cell or tissue sample from a subject (e.g. test subject) or cell line that comprises cellular RNA, optionally tumour tissue/cells and/or 2) RNA derived from such a sample. For example, the sample can be a biopsy, including a needle aspirate, such as a fine needle aspirate, a core biopsy, a brush biopsy and/or a laparoscopic biopsy. The sample can, for
30 example, be a "post-treatment" sample wherein the sample is obtained after one or more cytotoxic, e.g. cancer treatments, or a "base-line sample" which is optionally pre-

treatment or taken at an earlier time-point than the post-treatment sample, and is for example, used as a base line for assessing or monitoring response to a cytotoxic treatment. The tumour tissue/cells can be any tissue or cells, for example cancerous tissue or cells, for example ovarian cancer, prostate cancer, lung cancer, sarcomas, leukemia, lymphoma or multiple myeloma or colon cancer and/or any cancer for example cancers treatable by an anthracycline and/or taxane, including any subtype thereof including, for example, HER2+/-, ER+/-, PR+/-, topoisomerase +/- or triple negative breast cancer.

[0035] The term "RNA disruption assay score" or "RDA score" as used herein is a score or measure indicative of the extent of RNA degradation and/or likelihood of response to treatment, assessed on the basis of an RNA disruption assay described herein. The RDA score can for example be the output of a calculation based on electropherogram features described herein, for example represented as an RNA Disruption Index (RDI), which is described below, or can be a transformed scale, defined by clinical RDA zones based on cut-offs or thresholds, wherein each score or RDA zone is associated with a likelihood of a response for example associated with a likelihood of responsive to a treatment, for example defined by NPV and/or PPV. In an embodiment, a high or increased RNA disruption assay score (for example compared to a baseline sample or control) is indicative of high or increased RNA disruption. For example in such an embodiment, the higher the RNA disruption score, the greater the RNA disruption (and a decrease in RNA integrity). Any scale can be employed, for example 3, 10, or 600. A person skilled in the art would recognize that the scale could also be inverted (for example, by dividing the score into 1, e.g. $1/\text{RNA score}$) such that the lower the RNA disruption score, the greater the RNA disruption.

[0036] The term "RDA zones" as used herein refers to clinical zones associated with treatment response outcome comprising a range of RNA disruption scores, for example RDI values. Each RDA zone is defined by one or two boundaries each boundary corresponding to a selected threshold (e.g. corresponding to a desired NPV or PPV). Subject RDA scores (optionally RDI values) that fall within the clinical zones that are associated with or are predictive of treatment response, for example pCR. In the Examples, 3 RNA disruption assay zones are used, RDA zone 1, RDA zone 2 and RDA

- 18 -

zone 3, defined by selected NPV and/ or PPVs. A person skilled in the art would readily realize that any number of zones can be used each with different selected thresholds.

[0037] The term RNA Disruption Index "RDI" as used herein is a value generated using RDA and can be a ratio of features defined herein of the output of LDA or quadratic discriminant analysis of features described herein. The RDI values determined from a group of known response outcome patients (e.g. that are determined to be associated with a particular NPV or PPV) can be used to define the thresholds boundary for RDA zones.

[0038] The term "RDA zone 1" as used herein refers to a range of RDA scores (e.g. RDI values) that have a negative predictive value (NPV) of at least 0.8, at least 0.85, at least 0.9, at least 0.95, at least 0.96, at least 0.97, at least 0.98 or greater. These numbers and the associated thresholds, are for example based on the pCR in the Ma-22 data set and RDA measurement at third cycle used in the Examples. In an example embodiment, RDA Zone 1 is equal to an RDI of equal or less than 10 calculated using the feature Intermediate Area/(28S+18S Areas). In other embodiments, other features or combination of features can be used. The "RDA zone 1" can be defined to include any set of scores by selecting the desired NPV.

[0039] The term "RDA zone 2" as used herein refers to a range of RDA scores (e.g. RDI values) falling between RDA zone 1 and 3, and can be considered an intermediate or indeterminate zone.

[0040] The term "RDA zone 3" as used herein refers to a range of RDA scores (e.g. RDI values) that have a positive predictive value (PPV) of at least 0.15, at least 0.16, at least 0.17, at least 0.18, at least 0.19, at least 0.2 or greater. These numbers and the associated thresholds are for example based on the pCR in the Ma-22 data set and RDA measurement at third cycle used in the Examples. The "RDA zone 3" can be defined to include any set of scores by selecting the desired PPV.

[0041] The term "response", "responding" or "responsive" as used herein refers to a cell such as a cancerous cell or tumor response to a cytotoxic treatment such as a chemotherapeutic and/or radiation treatment, where the cells for example cancer cells or a subset of cancer cells within a tumour respond to the treatment in terms of RNA

disruption – e.g. in the context of a subject, the cells show significant treatment induced RNA degradation, and the subject has a positive treatment outcome, for example, reduction of tumour burden, disease stabilization, improved survival such as disease free survival and/or pathological complete response. In an embodiment, a subject who responds to a cytotoxic treatment with for example, one or more of a reduction of tumour burden, disease stabilization, improved survival such as disease free survival and/or pathological complete response, is considered, a “Responder” and/or a subject who does not for example, exhibit reduction in one or more of tumour burden, disease stabilization, improved survival such as disease free survival and/or pathological complete response is considered a “Non-Responder”.

[0042] The term “resistant” as used herein in reference to a cell such as a cancerous cell or tumour, refers to a cell or tumour response to a cytotoxic treatment, where the cells, e.g. cancer cells or subset of cancer cells within the tumour show no or little response to the treatment in terms of RNA disruption and a negative treatment outcome, for example disease progression and/or a lack of treatment benefit for the subject having the cancerous tumour.

[0043] The term “decreased RNA concentration” as used herein means an RNA concentration that is at least 20%, at least 25%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80% or at least 90% less than a control or corresponding reference value for example a pretreatment sample or for example a maximal value. RNA concentration can be determined by a number of methods including for example microcapillary electrophoresis, for example using for example an Agilent Bioanalyzer machine, an Experion® Capillary Electrophoresis System with its equivalent RNA Quality Index (RQI), Nanodrop® (Thermo Scientific, Inc.) or other equivalent systems, such as those manufactured by Applied Biosystems, Lumex, or Beckman Coulter Corporation or similar system. The RNA concentration can be based on UV absorbance, for example by assessing UV absorbance at 260 nm.

[0044] The term “low risk” as used in relation to progression refers to risk sufficiently less than average risk (e.g. decreased probability) calculated for a group of patients for example with the same cancer, treated similarly such that the patients would have high probability of survival with conservative therapy; and high risk of progression

- 20 -

means greater than average risk (e.g. increased probability) compared to the low risk group of patients.

[0045] The term “stable RNA integrity” or “lack of RNA disruption” as used herein means RNA that is not degraded appreciably, for example as compared to an appropriate comparator sample or the expected RNA integrity for the cell type of tissue, for example, less than a 10% decrease in integrity or less than a 15% decrease in integrity.

[0046] The term “RNA isolating or stabilizing composition” as used herein refers to any composition that inhibits RNase activity sufficiently and/or stabilizes RNA preventing RNA degradation.

[0047] The term “RNA integrity” as used herein means the degree of intactness of the RNA following extraction or isolation from the cell or tissue e.g. whether the isolated RNA is degraded (e.g. disrupted and/or autolytically degraded). High RNA integrity is commonly taken as meaning little to no degradation, for example less than a 30%, less than 25%, less than 20% or less than 10% decrease from a maximal RNA integrity and/or control sample or the retention of capacity to amplify mRNAs of interest following extraction or isolation. Low integrity RNA is, for example, RNA that exhibits greater than 20%, greater than 25%, greater than 30%, greater than 35%, greater than 40%, greater than 45%, greater than 50%, greater than 55%, greater than 60%, greater than 65%, greater than 70%, or greater than 75% decrease from maximal and/or a control sample or decreased capacity to amplify mRNAs of interest when they are known to be present in controls in RNA following extraction or isolation. RNA integrity can also be represented as an absolute number. For example, one RNA integrity scale assigns a number from 1 to 600 wherein an increasing score is associated with greater RNA disruption/degradation. For example, a sample that comprises a score of 500 comprises more RNA degradation than a sample scoring 400.

[0048] As used herein, and as well understood in the art, “treatment” is an approach for obtaining beneficial or desired results, including clinical results. Beneficial or desired clinical results can include, but are not limited to, alleviation or amelioration of one or more symptoms or conditions, diminishment of extent of disease, stabilized (i.e. not worsening) state of disease, preventing spread of disease, delay or slowing of

- 21 -

disease progression, reversal of disease, amelioration or palliation of the disease state, and remission (whether partial or total), whether detectable or undetectable. "Treatment" can also mean prolonging survival as compared to expected survival if not receiving treatment.

5 **[0049]** The term "signal induced cytotoxic treatment" as used herein means any agent or radiation that can induce cell death that is used in the treatment of disease for example cancer, including for example, traditional and non-traditional chemotherapy (e.g. targeted therapies), radiation treatment, hormonal treatment (e.g. for responsive cancers) and combinations thereof. Such agents include but are not limited to
10 microtubule stabilizing agents such as Docetaxel and paclitaxel, DNA synthesis inhibitors such Epirubicin, inhibitors of Her2 Receptor such as Trastuzumab, DNA cross-linking agents such as Mafosfamide, carboplatin and cisplatin, VEGFA inhibitors such as Bevacizumab, Receptor Tyrosine Kinase inhibitors such as Sunitinib and Toceranib, Bisphosphonates such as Zoledronic acid, Thymidylate synthase inhibitors such as 5-
15 fluorouracil.

[0050] The term "radiation" in relation to a treatment means any energy, photon or particle, applied to a tumour, including for example ionizing radiation.

[0051] The term "dose" as used herein in reference to radiation refers to an individual radiation exposure either administered at each time within a schedule or the
20 total amount of radiation exposure within a schedule. With respect to a chemotherapy treatment, a dose means an amount of an individual drug either administered at each time within a schedule or the total amount of each drug administered within a schedule or the total amount of drug administered during a course of chemotherapy.

[0052] The term "changing cancer treatment" or "altering cancer treatment" as
25 used herein includes for example one or more of changing the dosage level of the radiation, discontinuing the treatment, adding a chemotherapeutic agent(s) to the treatment or changing to an alternate cancer treatment such as a drug therapy or surgery.

[0053] In understanding the scope of the present disclosure, the term
30 "comprising" and its derivatives, as used herein, are intended to be open ended terms

- 22 -

that specify the presence of the stated features, elements, components, groups, integers, and/or steps, but do not exclude the presence of other unstated features, elements, components, groups, integers and/or steps. The foregoing also applies to words having similar meanings such as the terms, "including", "having" and their
5 derivatives.

[0054] In understanding the scope of the present disclosure, terms of degree such as "substantially", "about" and "approximately" as used herein mean a reasonable amount of deviation of the modified term such that the end result is not significantly changed. These terms of degree should be construed as including a deviation of at least
10 $\pm 5\%$ of the modified term if this deviation would not negate the meaning of the word it modifies.

[0055] In understanding the scope of the present disclosure, the term "consisting" and its derivatives, as used herein, are intended to be close ended terms that specify the presence of stated features, elements, components, groups, integers,
15 and/or steps, and also exclude the presence of other unstated features, elements, components, groups, integers and/or steps.

[0056] The recitation of numerical ranges by endpoints herein includes all numbers and fractions subsumed within that range (e.g. 1 to 5 includes 1, 1.5, 2, 2.75, 3, 3.90, 4, and 5). It is also to be understood that all numbers and fractions thereof are
20 presumed to be modified by the term "about." Further, it is to be understood that "a," "an," and "the" include plural referents unless the content clearly dictates otherwise. The term "about" means plus or minus 0.1 to 50%, 5-50%, or 10-40%, preferably 10-20%, more preferably 10% or 15%, of the number to which reference is being made.

[0057] Further, the definitions and embodiments described in particular sections
25 are intended to be applicable to other embodiments herein described for which they are suitable as would be understood by a person skilled in the art. For example, in the following passages, different aspects of the invention are defined in more detail. Each aspect so defined may be combined with any other aspect or aspects unless clearly indicated to the contrary. In particular, any feature indicated as being preferred or
30 advantageous may be combined with any other feature or features indicated as being preferred or advantageous.

II. Assays, Methods and Products

[0058] PCT application PCT/CA2008/001561 filed September 5th, 2008, incorporated herein by reference in its entirety teaches, for example, a method of
5 monitoring response to chemotherapy in patients with locally advanced breast cancer by monitoring the ability of the chemotherapy agents to induce degradation of RNA as measured, for example, by the RNA Integrity Value (RIN) (Schroeder, Mueller et al. 2006) as well as the ratio of 28S:18S ribosomal RNA. Using data from a national clinical trial (CAN-NCIC-CTG-MA.22; hereinafter "MA22"), it was demonstrated that tumour RIN
10 values fell significantly in some breast cancer patients upon treatment with epirubicin/docetaxel chemotherapy and that low tumour RIN values mid-treatment could be significantly correlated with high drug dose levels for the regimen ($p=0.05$) (Parissenti, Chapman et al. 2010). They also were able to show a correlation between low mid-treatment tumour RIN values and a pathologic complete response within the
15 breast and axilla post-treatment.

[0059] The MA22 RNA electropherogram data was examined to determine components of the electropherogram correlating with clinical response. It became evident that the Agilent Bioanalyzer software misidentified either the 28S peak or the 18S peak on electropherograms in >20% of the samples. For example, in at least some
20 of these samples, a small portion of the tip of a peak was identified as the whole peak. However, The Agilent 2100 Bioanalyzer software (Agilent Expert) (Mueller 2004; Vespucci 2005) was able to correctly identify the 28S and 18S peaks on electropherograms of intact RNA which are identifiable by eye. However, in samples with disrupted RNA, electropherogram peaks were misidentified by Agilent RIN software
25 although the correct peaks remained in appropriate locations visually on the electropherogram. The Agilent method which generates RIN values, hereafter referred to as the RIN algorithm, was based on RNA samples which were either intact or had undergone complete autolytic degradation and had few "partially degraded" samples (Schroeder, Mueller et al. 2006). Methods such as the Agilent method are generally for
30 the purpose of assessing autolytic degradation.

- 24 -

[0060] Importantly, the current RIN algorithm's misidentification of the 28S and 18S peaks resulted in aberrant 28S:18S ratios (often having a "0" value) or RIN values that were designated N/A (see for example **FIGS. 18A-18F**).

5 **[0061]** It is disclosed herein that cytotoxic treatment such as chemotherapy and/or radiation therapy induced RNA disruption comprises different profile features from autolytically degraded RNA. For example, **FIG. 1** shows an electropherogram trace of a tumour sample in saline soaked gauze for 24 hours where autolytic degradation takes place. A large peak of small fragments is visible at approximately the 26 second mark of the trace. There is also a loss of the 28S peak.

10 **[0062]** Cells treated with chemotherapeutic drugs show a different pattern of RNA disruption. For example, **FIG. 2** shows human ovarian cancer cell line A2780 treated with chemotherapeutic docetaxel and shows an intermediate peak and a peak adjacent to the 18S peak. **FIG. 3** shows A2780 cells treated with chemotherapeutic epirubicin. The electropherogram shows intermediate bands between the 18S and 28S
15 peaks. **FIG. 9D** shows an image of a gel of RNA from A2780 cells treated with increasing concentrations of chemotherapeutic docetaxel. Additional bands are visible in cells treated with 0.05 uM docetaxel. **FIG. 9E** shown an electropherogram of the 0.005 and 0.2 uM time-points with the x axis being the size of the nucleotides. The electropherogram of the 0.2 uM time-point shows intermediate bands between the 18S
20 and 28S peaks. In addition, the 28S and 18S peaks are shifted relative to those of **FIG. 1** which shows autolytic RNA degradation.

[0063] It was also identified that the peaks misidentified using the RIN algorithm were found in RNA samples which had large "intermediate banding" peaks located between the 28S peak and 18S peak (see the examples in **FIGS. 18A-F** and **FIG. 19**).

25 **[0064]** A similar pattern is seen with radiation treatment. Additional bands are also visible in cells treated with radiation.

[0065] Accordingly, an aspect of the disclosure includes assays, methods and apparatuses for determining if a cancer patient is responsive to a chemotherapeutic drug. Generally, the method comprises determining features that are obtained from at
30 least two of four different shifted regions of the electropherogram due to cytotoxic

- 25 -

treatment: the shifted 28S peak, the shifted 18S peak, at least one peak in the "intermediate banding" region where the intermediate banding region is between the shifted 18S peak and the shifted 28S peak, and at least one peak in the "low banding" region where the low banding region is below the shifted 18S peak. For example, a scale of 0 to 100 units (the scale is described in more detail with regards to **FIG. 5B**) may be selected which includes all RNA migrating between 9.5 and 42.8 seconds under the condition of a Caliper Nano chip run on an Agilent 2100 Bioanalyzer.

[0066] Referring now to **FIG. 4**, shown therein is a block diagram of an example embodiment of an apparatus 100 that can be used to assess RNA disruption. The apparatus 100 is configured to perform a method to assess the electropherogram of RNA samples such as cancer RNA samples after the cancer samples have been subjected to a chemotherapeutic drug and/or radiation. The assessment results in a RNA Disruption Assay (RDA) score that incorporates various features based on the electropherogram data as described herein. A high RDA score is associated, for example, with positive treatment response such as subsequent pCR, while a low RDA score indicates that patients are unlikely to receive long term chemotherapy and/or radiation benefit. The methods described herein detect peaks and abnormal peaks in the electropherogram with increased resolution and increased robustness to help assess the effectiveness of chemotherapy and/or radiation treatment. Accordingly, the application of the method described herein depends on accurate recording of fluorescent intensities during the separation of RNA molecules but is independent of commercial software used to analyze peaks in the electropherogram data.

[0067] In general, the apparatus 100 comprises at least one processing unit 102, a display 104, a user interface 106, data interface 108, Input/Output (I/O) hardware 110, a wireless module 112, a power source 114, a memory 116, and an electropherogram device 118. The memory 116 comprises software code for implementing an operating system 120, a file system 122, various programs 124 including an RNA disruption analysis module 126 and at least one database 128. The apparatus 100 can be a standalone device or can incorporate a desktop computer, a laptop, a mobile device, a smart phone, a cell phone, a tablet, a personal digital assistant, and the like.

- 26 -

[0068] There can be various alternative embodiments for the apparatus 100 in which one or more of the components are not used. For example, in some embodiments the wireless module 112 is optional. Some embodiments may use either the wireless module 112 or the data interface 108. Furthermore, in some embodiments, the electropherogram device 118 is not included as the electropherogram data may be obtained and sent, for example, via a data communication network or wireless communication, or otherwise provided to the apparatus 100, for example, via a CD, a USB drive, a USB key and the like. In these cases, the electropherogram data is typically available in a data file that can be in a spreadsheet format or other numeric format that provides a time series of fluorescence intensity for one or more samples. In these cases, a medical professional, medical institution or a research institution can obtain the electropherogram data and provide this data for analysis by the apparatus 100. For example, there can be situations in which a second institution has the apparatus 100, receives the electropherogram data, analyzes the electropherogram data using the apparatus 100 and then sends the results to the entity which originally provided the electropherogram data.

[0069] The processing unit 102 controls the operation of the apparatus 100 and can be any suitable processor depending on the configuration of the apparatus 100 as is known by those skilled in the art. In some embodiments, the processing unit 102 can be implemented using an Application Specific Integrated Circuit (ASIC), a programmable logic array, or a collection of discrete analog and discrete circuits.

[0070] The display 104 can be any suitable display that provides visual information depending on the configuration of the apparatus 100. For instance, the display 104 can be a cathode ray tube, flat-screen monitor and the like if the apparatus 100 is a computer. In other cases, the display 104 can be a display suitable for a laptop, tablet or handheld device such as an LCD-based display and the like.

[0071] The user interface 106 can include at least one of a graphical user interface, a mouse, a keyboard, a touch screen, a thumbwheel, a track-pad, a track-ball, a card-reader, voice recognition software and the like again depending on the particular implementation of the apparatus 100. In some cases, some of these components can be integrated with one another.

- 27 -

[0072] The data interface 108 can be any interface that allows the apparatus 100 to communicate with other devices or computers. In some cases, the data interface 108 can include at least one of a serial port, a parallel port, or a USB port that provides USB connectivity. The data interface 108 can also include at least one of an Internet or local
5 area network connection through Ethernet, Firewire or modem connections or through a digital subscriber line. Various combinations of these elements can be incorporated within the data interface 108.

[0073] The I/O hardware 110 can include at least one of a microphone, a CD-ROM drive, a CD-ROM read/write drive, a speaker and a printer. The wireless module
10 112 can be a radio that communicates utilizing CDMA, GSM, GPRS or Bluetooth protocol according to appropriate standards such as IEEE 802.11a, 802.11b, or 802.11g. The power source 114 can be any suitable power source that provides power to the apparatus 100 such as a power adaptor or a rechargeable battery pack depending on the implementation of the apparatus 100 as is known by those skilled in the art.

[0074] The memory 116 can include RAM and flash memory elements as well as other storage elements such as disk drives and hard drives. The memory 116 is used to store an operating system 120, a file system 122 and programs 124 as is commonly known by those skilled in the art. For instance, the operating system 120 and the file system 122 provide various basic operational processes for the apparatus 100. The
20 programs 124 include various user programs so that a user can interact with the apparatus 100 including viewing, analyzing and reporting data as well as possibly sending and receiving messages depending on the implementation of the apparatus 100.

[0075] The memory 116 also stores an RNA Disruption analysis module 126 and
25 one or more databases 128. The RNA Disruption analysis module 126 can determine RNA Disruption values for samples obtained from a patient that is receiving a form of chemotherapy and/or radiation according to various features and methods that are described in more detail with respect to **FIGS. 5A**, and **5B** or **12A**. In some embodiments, the RNA Disruption analysis module 126 can be implemented generally
30 by the processing unit 102 which can be a processor that is programmed with firmware

- 28 -

or that runs executable code, or is implemented by discrete circuit components, depending on the implementation of the apparatus 100.

[0076] In general, the processing unit 102 is configured to determine values for features from at least two shifted regions of the at least one electropherogram dataset and to determine an RDA score based on a combination of the values of the features. Accordingly, the processing unit 102, in its various alternative implementations, can act as means for determining values for features from at least two shifted regions of the at least one electropherogram dataset and means for determining an RDA score based on a combination of the values of the features. The shifting is due to the treatment by the cytotoxic treatment e.g. chemotherapy and/or radiation.

[0077] The processing unit 102 is coupled to a data input to receive at least one electropherogram dataset corresponding to a unique biological sample comprising cellular RNA, such as cancerous tumour RNA, at various time points including at least one of during or after the treatment. In some cases, the electropherogram dataset can correspond to (i.e. is derived from) a unique biological sample that has not undergone treatment. In some embodiments, at least one of the electropherogram device 18, the data interface 108, the wireless module 112, the I/O hardware 110, the memory 116 and other suitable devices can act as the data input or as means for obtaining at least one electropherogram dataset. The processing unit 102 is also coupled to a data output to convey an indication of the RDA score such as an RDI score, an RDA zone, or any other relevant analytical RNA data for at least one subject. In some embodiments, at least one of the display 104, wireless module 112, data interface 108, memory 116, and other suitable devices can act as the data output.

[0078] The databases 128 can be used to store data for samples that are assessed with the RNA Disruption analysis module 126. The databases 128 can also store other information required for the operation of the programs 124 or the operating system 120 such as dynamically linked libraries and the like.

[0079] The electropherogram device 118 is configured to produce an electropherogram of a sample. In general, the electropherogram device 118 employs a gel having a certain porosity in order to fractionate the RNA of the sample by size and charge of the particles. Fluorescence information related to the fractionated RNA is then

- 29 -

obtained to produce an electropherogram trace for the sample. The electropherogram trace is then digitized to form an electropherogram dataset. The electropherogram device 118 can be implemented in a variety of ways. For example, in some embodiments, the electropherogram device 118 is an Agilent 2100 BioAnalyzer that
5 employs Agilent RNA 6000 Nano kits and Caliper Technology's RNA Nanochips.

[0080] The apparatus 100 comprises at least one user interface and the processing unit 102 communicates with at least one of these user interfaces to receive electropherogram data as well as other information such as information on the individual from which the electropherogram was derived. This data can be received through the
10 data interface 108, the I/O hardware 110, the wireless module 112, the memory 116 or the electropherogram device 118 depending on the particular implementation of the apparatus 100 and the type of data. The processing unit 102 can communicate with either one of these elements as well as the display 104 or the I/O hardware 110 in order to output the RDA score. For instance, the apparatus 100 can output the RDA score to a
15 user of the apparatus 100. In addition, users of the apparatus 100 can communicate the resulting RDA score via a network connection to a remote system for storage and/or further analysis by other medical personnel. This communication can also include email communication in some embodiments.

[0081] In an alternative embodiment, the apparatus 100 can be a computer that
20 acts as a web server and provides content for a web site. One of the webpages on the website can be a webpage that provides one or more RDA scores from samples as described herein. In this case, a user can interact with the webpage to provide the electropherogram data (e.g. by uploading the data) and the apparatus 100 can analyze the electropherogram data and display the RDA score on the web page and/or email the
25 RDA score to the user and/or provide the RDA score for download by the user. Alternatively, the electropherogram data can be sent to the web server in an email communication, and the web server can analyze the electropherogram data to determine the RDA score and then send this information to a predefined destination, via email communication.

30 **[0082]** Referring now to **FIG. 5A**, shown therein is a flowchart of an example embodiment of a method 200 that can be employed to assess RNA disruption. The

- 30 -

method 200 is generally performed by the processing unit 102 in concert with additional elements of the apparatus 100, when required, in order to determine an RNA disruption assay score for electropherogram data that has been obtained from at least one sample at the same point in time.

5 **[0083]** Accordingly, an aspect of the present disclosure includes a method for performing an RNA Disruption Assay (RDA) for cellular RNA in response to a cytotoxic treatment, wherein the method comprises:

obtaining at least one electropherogram dataset corresponding to a unique biological sample comprising cellular RNA at a time point, during the course of
10 treatment or after completion of the treatment;

determining values for features from at least two shifted regions of the at least one electropherogram dataset, the shifting being due to the treatment; and

determining an RDA score based on a combination of the values of the features.

15 **[0084]** At 202, electropherogram data is generated for a sample. This may include generating several electropherogram datasets each from a unique biological sample that is taken from a patient or test subject (or if in vitro harvested) at a certain point in time, such as during or after cytotoxic treatment, for example, which may also be referred to as pre-therapy, mid-therapy and post-therapy respectively (also pre-
20 treatment, mid-treatment and post-treatment) where therapy is defined as a series of chemotherapeutic and/or radiation administrations (e.g. injections, exposures) that are performed over a certain period of time such as minutes, days or weeks. In general, an electropherogram dataset is a digitized version of electropherogram traces that are generated for a sample. The electropherogram dataset is an array of numbers that are
25 processed to generate an RDA score. In some embodiments, the electropherogram dataset can be in a spreadsheet format such as a Comma Value Separated (CSV) data file or an EXCEL data file format. In some embodiments, 202 of method 200 can be optional if the electropherogram data is produced by another entity and is then sent to another location for analysis at which case the method 200 starts at 204.

[0085] At 204, at least one electropherogram dataset(s) is obtained for a unique biological sample. The number of dataset(s) that are obtained depend on the features that are used to determine the RDA score. For example, some features can be used to determine the RDA score that require values from only one electropherogram dataset
5 generated from one sample. However, some features can be used to determine the RDA score that require several electropherogram datasets generated from several unique biological samples, such as two or three datasets obtained from two or three unique biological samples, for example. In either case, the electropherogram dataset is typically obtained from a storage device such as a memory element 116, a CD-ROM, a USB key,
10 a portable hard drive, from a database, from a network element, or from data that is uploaded over the Internet.

[0086] At 206, values are determined for features using two identifying ranges to allow for potential shifting of at least one of the 18S peak and 28S peak in the electropherogram dataset(s). For example, the identifying regions may be at least two
15 shifted regions of the electropherogram dataset(s). The shifted regions that are analyzed in order to determine the values for the features include at least two of a shifted 18S region, a shifted 28S region, an intermediate banding region and a low banding region. The intermediate banding region is defined as the region between the 28S and 18S peaks. The low banding region includes all banding below the 18S peak but does not
20 include the marker region. **FIG. 5C** is an example electropherogram in which these various peaks and regions are illustrated. The 28S and 18S peaks are defined to be within specific x axis (i.e. the time axis) ranges and can contain multiple maxima. However, an alternative method can be used for the identification of the 18S and 28S peaks and is described below in more detail with respect to **FIG. 12A**. In this case, the
25 identifying ranges are two broad regions that are defined (e.g. having a certain size) in order to locate the 18S and 28S peaks regardless of whether these peaks are shifted or otherwise distributed in the electropherogram dataset(s). Regardless of whether the method of **FIG. 5B** or the method of **FIG. 12A** is used, various features can be defined based on these regions and/or peaks. For example, the features can include one or
30 more of the area of an 18S peak, the area of a 28S peak, the area of an intermediate banding region, the area of a low banding region, the area of one or more band sub-regions, and the total area. The total area is the sum of the area of the 18S peak, the

- 32 -

28S peak, the intermediate banding region and the low banding region. Other features can include one or more of the width of the 18S peak and the width of the 28S peak. The area of these peaks and regions can be calculated in general using well-known mathematical techniques such as the trapezoidal rule for numerical integration
5 (Atkinson, 1989). However, the starting points of these peaks and regions are different due to the chemotherapeutic effect on the samples. Accordingly, in some cases, additional rules have to be used, in order to correctly determine the values of these features.

[0087] Another example embodiment of an improved method for the
10 identification of the 18S and 28S peaks and the calculation of certain features associated with these peaks or neighboring regions that can be used in act 206 of method 200 is described with respect to **FIG. 12A**. The low banding region can also be divided into the "low C", "low B" and "low A" sub-regions or banding regions as demonstrated in **FIG. 12B** for use in some methods described herein. It has been found
15 that the low A banding region may contain RNA due to autolytic degradation as well as due to other effects and that the low C banding region is an important region in assessing the effect on RNA due to various external stimuli such as cytotoxic treatments, for example. It has also been found that in response to certain drugs, the RNA starts to spread to the low C banding region, then to the low B banding region and
20 then to the low A banding region.

[0088] At 208, an RDA score is determined based on a combination of the values for various features. These combinations can include the ratios of some of these features. In some cases, only one electropherogram dataset from one sample is used to determine an RDA score. In other cases, several electropherogram datasets from
25 several unique biological samples obtained at the same time are used to determine an RDA score. In these latter cases, the values can be combined using a mathematical function such as the minimum function or the maximum function.

[0089] Features can also be based on other components of the shifted regions of the electropherogram such as one or more of a peak height, a peak width and a peak
30 position. For example, in one embodiment, it may be possible for the combination of

- 33 -

features to be based on the 18S and 28S peak widths to discriminate between Responders and Non-Responders.

[0090] In one embodiment, the combination of features comprises a ratio of the 28S peak area in the shifted 28S region to the 18S peak area in the shifted 18S region, which can be represented by the expressions "28S/18S" or "28S:18S".

[0091] In one embodiment, the combination of features comprises a ratio of the 28S peak area in the shifted 28S region to the total area, which can be represented by the expression "28S/total area".

[0092] In one embodiment, the combination of features comprises a ratio of the 18S peak area in the shifted 18S region to the total area, which can be represented by the expression "18S/total area".

[0093] In one embodiment, the combination of features comprises a ratio of the intermediate banding region area to the total area, which can be represented by the expression "intermediate area/total area".

[0094] In one embodiment, the combination of features comprises a ratio of the low banding region area to total area, which can be represented by the expression "low banding area/total area".

[0095] In one embodiment, the combination of features comprises a ratio of the intermediate banding region area to the addition of the 18S peak area in the shifted 18S region and the 28S peak area in the shifted 28S region, which can be represented by the expression "intermediate banding area/(18S+28S)".

[0096] In one embodiment, the combination of features comprises a ratio of the low banding region area to the addition of the 18S peak area in the shifted 18S region and the 28S peak area in the shifted 28S region, which can be represented by the expression "low banding area/(18S+28S)".

[0097] In one embodiment, the combination of features comprises a ratio of the low C banding region area to the sum of the 18S peak area in the shifted 18S region and the 28S peak area in the shifted 28S region, which can be represented by the expression "low C/(18S+28S)". It should be noted that 18S and 28S can be reversed in

- 34 -

any of these expression such as 28S+18S and it would not change the meaning of the expression.

[0098] In one embodiment, the combination of features comprises a ratio of the intermediate banding region area to the 18S peak area in the shifted 18S region, which
5 can be represented by the expression "intermediate banding area/18S".

[0099] In one embodiment, the combination of features comprises a ratio of the low banding region area to the 18S peak area in the shifted 18S region, which can be represented by the expression "low banding area/18S".

[00100] In one embodiment, the combination of features comprises a ratio of the
10 intermediate banding region area to the 28S peak area in the shifted 28S region, which can be represented by the expression "intermediate banding area/28S".

[00101] In one embodiment, the combination of features comprises a ratio of the low banding region area to the 28S peak area in the shifted 28S region, which can be represented by the expression "low banding area/28S".

[00102] In one embodiment, the combination of features comprises a ratio of the
15 addition of low banding region area and intermediate banding region area to the addition of the 18S peak area in the shifted 18S region and the 28S peak area in the shifted 28S region, which can be represented by the expression "(low banding area + intermediate banding area)/(18S + 28S)".

[00103] In one embodiment, the combination of features comprises a maximum
20 ratio determined from at least two electropherogram datasets. The ratio is defined as the intermediate banding region area divided by a sum of the 18S peak area in the shifted 18S region and the 28S peak area in the shifted 28S region. This can be represented by the expression "maximum intermediate banding area/(18S+28S)". In this case, as with
25 other calculations that use the maximum function with particular features that follow below, the ratios of intermediate banding area/(18S+28S) are calculated for two or more electropherogram datasets (derived from two or more unique biological samples at the same time point), the maximum of these ratios is calculated and then the logarithm is calculated. For example, two or more electropherogram datasets can be derived from

two or more biopsies taken at a time point and/or a biopsy that is divided into two or more samples.

[00104] In one embodiment, the combination of features comprises a minimum ratio determined from at least two electropherogram datasets. The ratio is defined as the intermediate banding region area divided by a sum of the 18S peak area in the shifted 18S region and the 28S peak area in the shifted 28S region. This can be represented by the expression "minimum intermediate banding area/(18S+28S)". In this case, as with other calculations that use the minimum function with particular features that follow below, the ratios of intermediate banding area/(18S+28S) are calculated for two or more electropherogram datasets (derived from two or more unique biological samples at the same time point) and the minimum of these ratios are is calculated.

[00105] In one embodiment, the combination of features comprises a maximum ratio determined from at least two electropherogram datasets. The ratio is defined as the intermediate banding region area divided by the total area. This can be represented by the expression "maximum intermediate banding area/total area".

[00106] In one embodiment, the combination of features comprises a minimum ratio determined from at least two electropherogram datasets. The ratio is defined as the low banding region area divided by the total area. This can be represented by the expression "minimum low banding area/total area".

[00107] In one embodiment, the combination of features comprises a ratio of low banding region area divided by the total area and a maximum of a second ratio determined from at least two electropherogram datasets, the second ratio being the intermediate banding area divided by the total area. These combinations of features can be represented by the expressions "low banding area/total area" and "maximum intermediate banding area/total area", respectively.

[00108] In one embodiment, the combination of features comprises a maximum ratio determined from at least two electropherogram datasets. The ratio is defined as the 28S peak area divided by the total area. This can be represented by the expression "maximum 28S/total area".

[00109] In one embodiment, the combination of features comprises a minimum ratio determined from at least two electropherogram datasets. The ratio is defined as the 28S peak area divided by the total area. This can be represented by the expression "minimum 28S/total area".

- 5 **[00110]** In one embodiment, the combination of features comprises a maximum ratio determined from at least two electropherogram datasets. The ratio is defined as the 18S peak area divided by the total area. This can be represented by the expression "maximum 18S/total area".

- 10 **[00111]** In one embodiment, the combination of features comprises a minimum ratio determined from at least two electropherogram datasets. The ratio is defined as the 18S peak area divided by the total area. This can be represented by the expression "minimum 18S/total area".

- 15 **[00112]** In one embodiment, the combination of features comprises a maximum ratio determined from at least two electropherogram datasets. The ratio is defined as the intermediate area divided by the 28S peak area in the shifted 28S region. This can be represented by the expression "maximum intermediate/28S".

- 20 **[00113]** In one embodiment, the combination of features comprises a minimum ratio determined from at least two electropherogram datasets. The ratio is defined as the intermediate area divided by the 28S peak area in the shifted 28S region. This can be represented by the expression "minimum intermediate/28S area".

[00114] It should be noted that in each of these feature combinations described above, either method 300 or method 400 (described in **FIG. 12A**) can be used to define the 18S and 28S peaks and the other regions. If method 400 is used rather than using shifted regions, broader regions are used to locate the 18S and 28S peaks.

- 25 **[00115]** At 210, the RDA score can be compared to a threshold to determine the effectiveness of the treatment. The threshold can be a scalar value determined based on conducting ROC analysis on a database of various experimental results to identify responders which can be done by selecting a threshold value to accomplish a high Positive Predictive Value (PPV) and a low false positive rate. The analysis can also be
30 conducted to identify non-responders by selecting a threshold value so that the non-

- 37 -

responders can be identified with a high Negative Predictive Value (NPV) and a low false negative rate. Accordingly, the goal is to have a threshold for high PPV and a threshold with high NPV and to minimize the distance between the PPV and NPV threshold values. In some embodiments, several thresholds can be used to define
5 several RDA zones which relate to different types of expected treatment results such as increased likelihood of a complete response, a partial response and no response (see **FIG. 7A** and/or **FIG. 14A** for examples of the RDA zones).

[00116] In an alternative embodiment, rather than comparing the RDA score to a threshold value, the RDA score may be compared to a threshold curve which may be
10 linear or non-linear or a threshold three dimensional region. In this case, the threshold line, threshold curve or threshold plane can be determined using LDA or QDA as is described with respect to **FIGS. 15B-15D**.

[00117] Referring now to **FIG. 5B**, shown therein is a flowchart of an example embodiment of a method 300 that can be employed to determine features related to 18S
15 and 28S peaks in the 18S and 28S shifted regions respectively (another method is described with respect to **FIG. 12A**). While these peaks are distinguishable from other peaks and have certain sizes, one or more rules can be used to increase the accuracy of the determination of these peaks and their resulting areas. However, not all of the following rules need to be applied. For example, when there are strong 18S and 28S
20 peaks and little or no intermediate banding, then steps 310 and 312 may not have to be performed.

[00118] At 302, boundaries for the shifted regions of the electropherogram are defined due to the effects of chemotherapy, or some other treatment regimen or effect, on the sample in order to properly determine the 18S and 28S peaks. After iterating
25 through a number of experimental results, about 95% of the 18S peaks were found in the range of 39.5 to 43.5 seconds (i.e. shifted 18S region) and those of the 28S peaks were in the range of 47.5 to 53.5 seconds (i.e. shifted 28S region); this analysis can be repeated for other treatment regiments. Accordingly, the boundary conditions for the shifted 18S and 28S regions were changed from the range of 38.5 to 46.5 seconds and
30 the range of 43 to 53.5 seconds to the range of 39.5 to 43.5 seconds and the range of 47.5 to 53.5 seconds respectively. It should be noted that these ranges can change

- 38 -

depending on different processing techniques such as that described with respect to **FIG. 12A**, for example. If the time axis (e.g. x axis) of the electropherogram datasets is converted to a scale of 0 to 100 units, then the boundary conditions for the shifted 18S region become 40.9 to 50.4 and the boundary conditions for the shifted 28S region become 58.1 to 69.5. The conversion includes mapping the end of the marker region as the 0 unit and a time point in the area of about 60 to 65 seconds, such as 63 seconds, for example, as the 100 unit. This resulted in 18S and 28S peaks being defined by more stringent conditions.

[00119] Most peaks were identified correctly using these shifted 18S and 28S regions. This may involve searching for all of the peak candidates first, and then choosing the one with largest area as the 18S peak or the 28S peak. Candidate peaks whose width is not within about 0.6 to 7.6 units for the 18S peak and not within about 0.8 to 10.5 units for the 28S peak, respectively, may be excluded from the candidate peak set, so that they will not be chosen as the 18S peak or the 28S peak. It has been found that the probability of misidentifying the 28S peak is much larger than that of misidentifying the 18S peak. Accordingly, one may determine the 18S peak first. Also, when identifying the 28S peaks, those 28S peak candidates whose distance to the 18S peak is greater than a distance threshold will be excluded from the peak candidates set so that they will not be chosen as the 28S peak.

[00120] At 304, the peak area is determined for a peak, such as the 18S and the 28S peaks, for example. To make this determination, the idea of a local minimum (Stewart, 2008) can be used to define the peak area. The local minimum of a function is the smallest value that the function takes at a point within a given neighborhood. Once the peak of 18S or 28S is known, the two local minimums closest to the peak can be found. The point corresponding to the local minimum on the left side of the peak (within the shifted region) can be defined as the starting position of the peak, and the point corresponding to the local minimum on the right side of the peak (within the shifted region) can be defined as the ending position of the peak. The peak area is defined as the area between these two local minimums. Accordingly, a peak area for a given peak is defined by an area between two local minimum values on either side of the peak in the given shifted region where the peak has one maximum value.

[00121] However, the peak can also be defined to have more than one maximum value in certain cases as is described below. At 306, it is determined if the peak should have more than one maximum value (i.e. more than one peak) depending on the area of the peak. This adjustment is used since many of the errors which were found with the RIN method were due to multiple maxima within the 28S and 18S peaks (see **FIG. 18**). The existence of multiple peaks or multiple values in a peak can be determined in several ways. By analyzing experimental data, if it was found that the value of a peak area was smaller than a threshold value, then the peak was allowed to comprise more than one maximum value. In other words, if the calculated area doesn't match the corresponding peak, i.e. the peak area is very small when the peak is actually large, then the peak is defined to be multi-peaked. For example, the peak could have two maximum values. In one example embodiment, the threshold is 0.2 FU-seconds. This condition successfully corrected for most of the previously improperly detected examples of the RIN method.

[00122] At 308, boundaries of the shifted 18S and 28S regions are adjusted if needed depending on the areas of these peaks with respect to each other. For example, if an 18S peak area in the shifted 18S region divided by total area multiplied by 100 (which is represented by 18S%) is less than about 8 times the 28S peak area in the shifted 28S region divided by total area multiplied by 100 (which is represented by 28S%), the shifted 18S region is defined to be the range of 42.8 to 51.4 units (where the time axis is converted to a scale of 0 to 100), i.e. the bounds of the shifted 18S region were extended from 42.8 to 50.4 units to about 42.8 to 51.4 units. In a similar fashion, if the 28S peak area in the shifted 28S region divided by total area multiplied by 100 (i.e. 28S%) is less than about 8 times an 18S peak area in the shifted 18S region divided by total area multiplied by 100 (i.e. 18S%), the shifted 28S region is defined to be the range of about 52.3 to 69.5 units (where the time axis is converted to a scale of 0 to 100), i.e. the bounds of the shifted 28S region were extended from about 58.1 to 69.5 units to about 52.3 to 69.5 units.

[00123] At 310, the peak widths are adjusted if needed, i.e. if the widths are defined to be too large based on the position of the starting and ending position of a peak and a study of experimental results. Accordingly, after the peak widths of 18S and

- 40 -

28S were calculated they were subsequently adjusted if the width was outside a width range such that the width of the 18S peak should be within 0.6 to 7.6 units and the width of the 28S should be within 0.8 to 10.5 units when the time axis is converted to a scale of 0 to 100.

- 5 **[00124]** At 312, the distance between the peaks is adjusted if they are calculated such that they are too far apart to enhance the definition of the 18S and 28S peaks. Accordingly, a location of at least one of the 18S and 28S peaks is adjusted by picking a different candidate peak if the distance between the 18S and 28S peaks is greater than a distance threshold. For example, in some embodiments, the distance between an
10 ending position of an 18S peak and a starting position of a 28S peak is defined to be less than about 17.1 units apart.

- [00125]** Accordingly, another aspect of the disclosure is a method for performing an RNA Disruption Assay (RDA) for cellular RNA, wherein the method comprises: obtaining at least one electropherogram dataset corresponding to a unique biological
15 sample comprising the cellular RNA at a time point; defining an 18S peak and a 28S peak from the at least one electropherogram dataset; determining at least one parameter value for both the 18S peak and the 28S peak; and redefining at least one of the 18S peak and the 28S peak when required according to one or more rules applied to the at least one parameter value. In at least some cases, the method can further include
20 determining an 18S peak area and a 28S peak area; and determining an RDA score based on at least one of the 18S peak area and the 28S peak area.

[00126] In some embodiments, the at least one parameter value comprises at least one of a peak area, a peak width and a peak location.

- [00127]** The one or more rules comprises at least one of determining if at least
25 one of the 18S peak and the 28S peak have more than one maximum value based on peak area, determining if boundaries where the 18S peak and the 28S peak are located require adjustment depending on comparing the areas of the 28S peak and 18S peak to one another, adjusting widths of the 18S peak and the 28S peak to be in a certain width range, and adjusting a location of at least one of the 18S peak and the 28S peak if a
30 distance between the 18S peak and 28S peak is greater than a distance threshold. In some cases, the two rules of adjusting widths of the 18S peak and the 28S peak to be in

- 41 -

a certain width range, and adjusting a location of at least one of the 18S peak and the 28S peak if a distance between the 18S peak and 28S peak is greater than a distance threshold are optional.

5 **[00128]** Once the 18S and 28S peaks are defined according to the method 200, other values can be determined such as the intermediate area, low banding area, etc.

[00129] It should be noted that in each of these feature combinations described below, either method 300 or method 400 (described in **FIG. 12A**) can be used to define the 18S and 28S peaks and the other regions. If method 400 is used than rather than using shifted regions, broader regions are used to locate the 18S and 28S peaks.

10 **[00130]** In one embodiment, the combination of features comprises a ratio of the intermediate area to the sum of the 28S peak area in the shifted 28S region and the 18S peak area in the shifted 18S region, which can be represented by the expression "Intermediate/(28S+18S)".

15 **[00131]** In one embodiment, the combination of features comprises a ratio of the sum of the low C banding region area and intermediate area to the sum of the 28S peak area in the shifted 28S region and the 18S peak area in the shifted 18S region, which can be represented by the expression "Intermediate + low C/(28S+18S)".

20 **[00132]** In one embodiment, the combination of features comprises a ratio of the sum of the low C banding region area and the intermediate area to the sum of the 28S peak area in the shifted 28S region and the 18S peak area in the shifted 18S peak region, which can be represented by the expression "(low C + Intermediate)/(28S+18S)".

25 **[00133]** In one embodiment, the combination of features comprises a maximum ratio determined from at least two electropherogram datasets. The ratio is defined as the intermediate area divided by the sum of the 18S and 28S peak areas in the shifted 18S and 28S regions respectively. This can be represented by the expression "maximum intermediate/(28S+18S)".

[00134] In some embodiments, the apparatus 100 and the computer readable medium can be configured to implement methods 200 and 300 or methods 200 and 400 or methods 200, 300 and 400 where method 400 is described with respect to **FIG. 12A**.

- 42 -

[00135] The step of obtaining the at least one electropherogram dataset comprises in an embodiment: separating the cellular RNA by electrophoresis; detecting one or more species of the separated cellular RNA; plotting the electropherogram and obtaining the at least one electropherogram dataset from the electropherogram plot.

- 5 **[00136]** Electrophoresis for separating RNA is a commonly used technique. For example, gels containing fluorescent dyes can be employed to detect separated RNA species. Automated microfluidics based electrophoresis methods and systems such as the Experion automated electrophoresis system (Bio Rad Laboratories, Inc.) and the Agilent 2100 Bioanalyzer (Agilent Technologic, Inc.) have been developed which can be
10 used to separate RNA species. These systems use very small amounts of RNA and conduct microcapillary electrophoresis in channels of microchips.

[00137] Accordingly, the gel electrophoresis is microcapillary electrophoresis. In another embodiment, the microcapillary electrophoresis comprises using an RNA chip for electrophoretically separating the RNA.

- 15 **[00138]** Detection of the separated cellular RNA is performed, for example, using RNA dyes such as fluorescent dyes which bind the RNA. The RNA dyes can be added to the electrophoresis gel permitting detection of the RNA by detecting the dye signal.

- [00139]** Cellular RNA for use in the assay can be, for example, isolated from a biological sample (including any biological sample comprising cancer cells including for
20 example a sample comprising cells of a cell line, human tissue or animal tissue) using techniques known in the art. For example, cells and/or a biopsy can be re-suspended/lysed in a RNA isolating or stabilizing solution such as RNAzolTM(Sigma/Aldrich Co.), RNAlaterTM (Qiagen Laboratories), RNA StableTM (Biomatrix Co.), RNA Protect Cell ReagentTM(Qiagen Laboratories) and RNA isolated
25 using one or more commercially available kits (according to manufacturer's directions), including, but not limited to, RNeasyTM(Qiagen Laboratories), miRNeasyTM.(Qiagen Laboratories), RNAzolTM(Sigma/Aldrich Co.), Total RNA Isolation Kit (Norgen Biotek Co. and Agilent Technologies, Inc.), PurelinkTM (Invitrogen Labs),5 PRIME PerfectPureTM, and GeneJET (Thermo Scientific) kits.

[00140] As mentioned, changes in RNA disruption can be used to monitor response to cytotoxic treatment. For example, **FIGS. 6A-6K, 13, 14A-D, 15A-15D** and others demonstrate that RDA can be used to discriminate subjects receiving chemotherapy with an increased likelihood of pCR from non-responders (i.e. non-pCR
5 responders) as well identify patients in an intermediate RDA zone. Changes in RNA disruption are also seen with radiation (e.g. see **FIGS. 10, 11A and 11B**). Without wishing to be bound to theory, RNA disruption in a tumour cell may take place before cell death and/or before changes in cellularity, particularly during early stages of treatment. Accordingly, the RDA is suitable for assessing subject response to a cytotoxic
10 treatment, such as a chemotherapy treatment.

[00141] Another aspect of the disclosure includes an assay comprising:
quantifying the amount of RNA disruption in a biological sample comprising cellular tumour RNA at a time point during or after cytotoxic treatment such as, but not limited to, chemotherapeutic treatment, radiation treatment, and cytotoxic adjuvant
15 treatment, for example, using an RNA disruption assay (RDA) as described herein to determine an RDA score; and
determining if the amount of RNA disruption of the cellular RNA is increased or not increased.

[00142] As mentioned this can involve determining whether a subject falls within
20 an RDA zone. For example, **FIGS. 7A and 14A**, which are log-log plots, show that responders and non-responders can be separated by RDA zones.

[00143] In an embodiment, the RDA zone information may be transmitted to the subject, optionally through a medical professional.

[00144] The tumour cell RNA can be from, for example, a tumour cell line, or can
25 be from a tumour cell sample from a subject such as a biopsy.

[00145] In one embodiment, the tumour cell RNA assessed is from an untreated sample (e.g. untreated tumour cells in vitro) or is from a subject at a time point before treatment (e.g. pre-therapy).

[00146] Reports of tumour samples showing aberrant 28S/18S ratios have been
30 reported.

[00147] For example, abnormal 28S:18S ratios have been reported in untreated tumour tissue (Skrypina, et al. 2003). It has been suggested that this loss in 28S is due to preferential cleavage within GC-rich regions (Johnson, Sandler, Lalancette, Hauser, Diamond, & Krawetz, 2011).

- 5 **[00148]** In one embodiment, the cellular RNA is isolated from a normal cell. For example, the present method and assays can be used to assay normal cell response to cytotoxic drugs such as, but not limited to, chemotherapeutics, for example, to assess drug toxicity on non-cancer cells.

- 10 **[00149]** In one embodiment, the determining step comprises comparing the quantified amount of RNA disruption in the biological sample with a control such as a threshold or reference value. In an embodiment, the reference value corresponds to a pretreatment value.

[00150] In an embodiment, the increase is relative to a control or an expected pretreatment score.

- 15 **[00151]** A pretreatment sample comprising for example cancerous tumour cell RNA can comprise relatively little or no RNA disruption. For example, the RNA integrity will be stable. For example, untreated breast cancer tumour biopsies assessed for RNA disruption using the methods described herein are typically identified as falling within zone 1. RNA degradation can be scored in some embodiments without relation to a
20 comparator (e.g. using an absolute scale). In some embodiments, threshold values can be used. Threshold values for example, which define a zone can be used to classify the subject. Accordingly a determination of an RDA score that falls for example within zone 2 or 3 is indicative that the sample contained disrupted RNA even in the absence of a comparator sample on the basis of the expected untreated amount of RNA disruption.
25 Accordingly, the determining step can comprise comparing to a priori determined value or an expected value for the tumour type.

- [00152]** In another embodiment, the RNA disruption assay (RDA) is for measuring cancer treatment, e.g. chemotherapy and/or radiation, induced RNA disruption. Assays employing the RDA can be used to identify which subjects and/or tumours respond to a
30 chemotherapy and/or radiation treatment.

- 45 -

[00153] Accordingly in another embodiment, the assay further comprises identifying the tumour as responsive to a cancer treatment, optionally a chemotherapeutic and/or radiation treatment, if the amount of RNA disruption (e.g. the RDA score) is increased relative to the reference or threshold value.

- 5 **[00154]** In another embodiment, the biological sample corresponds to cells, such as cancerous tumour cells, treated in vitro or in vivo.

[00155] FIG. 9I shows that sensitive A2780 cells show RNA disruption whereas resistant A2780 cells do not.

- [00156]** Also provided in one embodiment described herein, is an in vitro assay
10 for determining sensitivity of a cancerous tumour cell or a normal cell to a cancer treatment such as a chemotherapeutic and/or radiation treatment, the assay comprising:
treating a cell or cell population with a chemotherapeutic and/or with radiation;

- quantifying the amount of RNA disruption in a biological sample
15 comprising cellular RNA at a time point during or after treating the cell or cell population with the chemotherapeutic and/or radiation using an RNA Disruption Assay (RDA) as described herein optionally to determine an RDA score; and

- comparing the quantified amount of RNA disruption with a threshold or reference value, and if the amount of RNA disruption is increased relative to the
20 reference value, identifying the cell as sensitive to the chemotherapeutic and/or radiation treatment and if not increased (e.g. if not increased significantly) identifying the cell as resistant to the treatment.

- [00157]** In one embodiment, the RNA disruption is expressed as an RDA score. The RDA score is reflection of the amount of RNA disruption in the sample, with
25 increased RDA score reflecting increased RNA disruption.

[00158] As mentioned, the assays can be used to determine the likelihood that a subject will or is responding to a cancer treatment and/or their risk of progression.

- [00159]** Accordingly a further embodiment comprises an assay for determining if a subject is responding to a cancer treatment comprising: assaying a biological sample
30 obtained from the subject during and/or after the subject has received a cancer

- 46 -

5 treatment, optionally chemotherapy and/or radiation therapy for the quantity of RNA disruption using an RNA Disruption Assay (RDA) performed according to a method described herein, wherein the subject is identified as responding to the treatment if the amount of RNA disruption (e.g. the RDA score) is increased relative to the threshold or reference value.

[00160] In one embodiment, the responsiveness is indicative of treatment outcome, and an increased amount of RNA disruption relative to a threshold or reference value (indicative of increased chemotherapy induced RNA disruption) is predictive of a positive treatment outcome and a lack of RNA disruption relative to a threshold or reference value (e.g. indicative of little or no chemotherapy induced disruption) is predictive of negative treatment response (e.g. lack of long-term benefit).

[00161] In another embodiment, the positive treatment outcome is pathologic complete response following treatment, partial response, reduced risk of disease progression, stable disease or increased overall survival.

15 **[00162]** In one embodiment, the cancerous tumour sample assessed is obtained from a subject during and/or after the treatment, and/or the time point is, during and/or after the treatment.

20 **[00163]** In one embodiment, the methods or assays described herein are used to monitor treatment response. For example, a subsequent sample is compared to a previous sample such as a baseline sample or value. In another embodiment, the assay and methods described herein are employed in a clinical trial to assess and/or monitor which subjects are responding to the treatment. For example, subjects can be assigned treatments optionally based on expression analysis.

25 **[00164]** Biopsies can be taken, during and/or post treatment, and the RDA can be used to predict response, such as, for example, to predict if a subject is likely to benefit from remaining on a treatment (e.g. is a responder) or is likely to benefit from switching treatments thereby monitoring treatment response, adequate dosage etc. For subjects not responding according to the RDA, a subset can be switched to a new treatment. The outcome of patients that were switched can be compared to those that remained on the treatment predicted to be ineffective according to RDA score.

30

[00165] Accordingly in an embodiment, the method or assay further comprises selecting non-responders and randomizing participation in two or more arms of a clinical trial.

[00166] In yet another embodiment, the method or assay further comprises
5 selecting responders and randomizing participation in two or more arms of a clinical trial.

[00167] In another aspect, the disclosure includes a method for selecting subjects for a multi-arm clinical trial, the method comprising:
predicting if the subject receiving a cytotoxic treatment such as
chemotherapeutic and/or radiation treatment is responding or not responding to the
10 treatment according to a method for determining an RDA score described herein;
randomizing non-responders to two or more treatment arms; and
optionally randomizing responders to two or more treatment arms.

[00168] A further aspect includes a method for predicting the efficacy of a
cytotoxic treatment, such as a chemotherapy and/or radiation treatment, comprising:
15 stratifying a subject group into at least 2 subgroups;
treating the subjects of each subgroup with the chemotherapy treatment
for a suitable time;
determining an RDA score for each subject according to the method
described herein; and
20 predicting the treatment to be efficacious if the subjects have an RDA
score above a threshold or reference value and predicting the treatment not to be
efficacious if the subjects have an RDA score below a threshold or reference value.

[00169] Also provided is a method for predicting the efficacy of a chemotherapy
treatment for a subject, comprising:
25 treating the subject with chemotherapy treatment for a suitable time; and
determining an RDA score according to the method described herein;
wherein a treatment is predicted to be efficacious for the subject if the subject has an
RDA score above a threshold or reference value and predicted not to be efficacious for
the subject if the subject has an RDA score below a threshold or reference value.

[00170] If a subject is identified as not responding to a cytotoxic treatment, the subject's treatment can be altered or changed. For example, the treatment dose administered can be increased, radiation and/or a chemotherapeutic agent can be added to the regimen or the treatment can be discontinued and/or a different cytotherapy regimen initiated. The methods and assays can also be used to monitor response to the altered treatment. Other treatment modalities can also be employed if a subject is not responding to a cytotoxic, e.g. chemotherapy treatment such as radiation treatment or surgery.

[00171] Cytotoxic treatments such as chemotherapy and radiation are often administered in cycles wherein the cycle can comprise administration of 1, 2 or more doses per period, for example, per day, per week, per two week period. Protocols have been established for different drugs and different cancers. A treatment regimen can consist of 1, 2, 3, 4, 5, 6 or more cycles. Accordingly, in an embodiment, the RNA sample is at a time point after a first cycle. In another embodiment, the RNA sample that is assessed is obtained from a subject mid-treatment, and/or the time point is mid-treatment (e.g. for a 6 cycle treatment regimen, the RNA sample assessed is obtained from a subject and/or the time point is after 2, 3, or 4 cycles).

[00172] In one embodiment, the biological sample is obtained after the subject has received a first cycle of cytotoxic treatment for example a first cycle of a chemotherapeutic or radiation treatment. In another embodiment, the biological sample is obtained from the subject after the subject has received 2, 3, 4 or more cycles of the treatment. In yet another embodiment, the biological sample is obtained from the subject after completion of the treatment regimen.

[00173] In an embodiment, the biological sample is a biopsy such as a fine needle aspirate. In another embodiment the biological sample comprises a population of cancer cells isolated from the subject.

[00174] More than one biological sample can be assessed. For example, where multiple samples are taken at a time point and/or at the same time, for example, 2 or more fine needle aspirate biopsies, each sample can be assessed separately and the congruity of the samples can be assessed to determine any heterogeneity in treatment response.

[00175] In an embodiment, the RDA score indicative of responsiveness and/or positive treatment outcome correlates to a decrease in RNA integrity (e.g. increase of RNA disruption) of at least by 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80% or at least 90%.

- 5 **[00176]** In an embodiment, the peaks are not shifted. However, the method is still applied according to the methods described herein.

[00177] In one embodiment, the methods and assays described herein are applied with non-tumour RNA samples, for example normal cells that are sensitive to a drug for example chemosensitive to a chemotherapeutic and/or radiosensitive to radiation therapy.

10 **[00178]** Without wishing to be bound by theory, it is believed that RNA disruption is a process that affects stable cells in an intermediate or non-acute response. Accordingly, the methods, assays, and products may be applicable to any cell type undergoing stimulation induced cytotoxicity e.g. RNA disruption.

- 15 **[00179]** It has been demonstrated that the total RNA of Jurkat T leukemia cell line was susceptible to chemical attack by doxorubicin (Fimognari et al 2009).

[00180] The assays can be applied to any tumour type and/or subtype. For example, the cancerous tumour can be a breast cancer, lung cancer, a sarcoma, prostate cancer, colon cancer or ovarian cancer including any subtype thereof. In certain
20 embodiments, the methods and assays are applied to blood cancers such as leukemias, multiple myeloma or lymphomas. In an embodiment, the cancer is a cancer treatable by an anthracycline and/or a taxane and/or radiation. For example, prostate cancer is a cancer that can be treated by docetaxel and lung cancer and sarcomas can also be treated by anthracyclines and/or taxanes. Anthracyclines include for example
25 doxorubicin, daunorubicin, epirubicin, idarubicin, valrubicin and mitoxantrone. Taxanes for example include paclitaxel, docetaxel, larotaxel, Abraxane, docoxahexaenoic acid-linked paclitaxel, paclitaxel polyglumex, Ortataxel, Genexol, liposomal-encapsulated paclitaxel, and paclitaxel in a Vitamin E emulsion.

- 50 -

[00181] As demonstrated for example in **FIG. 10**, cancer cells demonstrate RNA disruption in response to radiation treatment. **FIGS. 11A** and **11B** provide a description of analyses where patients received chemotherapy and radiation therapy.

[00182] In an embodiment, the subject or cells have received 1, 2, 3, 4, 5, 6 or
5 more doses of radiation. In an embodiment, the cells and/or subject are treated with a dose of a chemotherapeutic agent prior to or concurrent with radiation treatment.

[00183] The MA22 dataset for example, included samples with differing ER, PR, Her2, topoisomerase status. As demonstrated in the Examples section, the method and assays described herein are applicable independent of ER, PR, Her2, topoisomerase
10 status. The methods have been used to assess breast cancer samples that vary in ER, PR and Her2 status including triple negative samples.

[00184] In one embodiment, the cancerous tumour is selected from breast cancer, prostate cancer, colon cancer and ovarian cancer or any subtype thereof.

[00185] In an embodiment, the cancerous tumour is selected from leukemia,
15 lymphoma or multiple myeloma and/or any cancer treatable by an anthracycline and/or taxane.

[00186] In one embodiment, the cancerous tumour is breast cancer.

[00187] In another embodiment, the breast cancer is HER2+ positive breast cancer.

[00188] A further aspect includes a method of treating a subject with cancer in
20 need thereof comprising:

administering a cytotoxic treatment optionally a chemotherapeutic and/or radiation treatment;

determining if the subject is responding to the treatment according to a
25 method disclosed herein; and

continuing administration of the treatment if the subject is responding and/or discontinuing administration of the treatment if the subject is not responding.

[00189] In an embodiment, the cytotoxic treatment dose, e.g. chemotherapeutic dose, is increased, for example wherein a subject sample includes an intermediate level

of RNA disruption, for example RDA zone 2. As subjects who fall within RDA zone 3, are more likely to be pathological complete responders, a subject whose RDA score does not fall within RDA zone may benefit from increased treatment dose.

[00190] In an embodiment, the chemotherapeutic treatment is selected from
5 microtubule stabilizing agents such as Docetaxel and paclitaxel, DNA synthesis inhibitors such Epirubicin, inhibitors of Her2 Receptor such as Trastuzumab, DNA cross-linking agents such as Mafosfamide, carboplatin and cisplatin, VEGFA inhibitors such as Bevacizumab, Receptor Tyrosine Kinase inhibitors such as Sunitinib and Toceranib, Bisphosphonates such as Zoledronic acid, and Thymidylate synthase inhibitors such as
10 5-fluorouracil.

[00191] In one embodiment, the chemotherapeutic treatment is selected from taxanes, anthracyclines, and vinca alkaloids such as vinblastine, alkylating agents such as cisplatin and nucleoside analogs such as 5-FU and combinations thereof. In an
15 *embodiment, the chemotherapeutic treatment comprises a drug selected from docetaxel, paclitaxel, epirubicin, doxorubicin, vinblastine and cisplatin and/or combinations thereof.* In an embodiment, the taxane is selected from paclitaxel, docetaxel, larotaxel, Abraxane, docoxahexaenoic acid-linked paclitaxel, paclitaxel polyglumex, Ortataxel, Genexol, liposomal-encapsulated paclitaxel, and paclitaxel in a Vitamin E emulsion.

[00192] Breast cancer can be treated, for example, with taxanes, anthracyclines
20 or, for example, paclitaxel, docetaxel, docetaxel and epirubicin and combinations thereof.

[00193] Ovarian cancer can be treated, for example, with taxanes, anthracyclines, and vinca alkaloids such as vinblastine, alkylating agents such as cisplatin and nucleoside analogs such as 5-FU and combinations thereof.

25 **[00194]** In an embodiment, the RNA integrity is determined for total RNA. In another embodiment, the RNA integrity is determined for ribosomal RNA. Ribosomal RNA includes, for example, 28S, 18S, 5.8S and 5S rRNA.

[00195] The methods, assays and products described herein can also be used to assess autolytic degradation. For example, autolytic degradation can be assessed by
30 determining the Area of LowA/Total Area. Measurement of this feature using xenograft

- 52 -

tumours incubated with saline for increasing lengths of time, demonstrates that autolytic degradation can be measured more often and may be measured more accurately than what is possible using BioAgilent's RIN software. For example, using the ratio of Low A Area/Total Area, a more precise mapping of the extent of autolytic degradation may be measured.

[00196] At least some of the elements of the methods that are described herein that are implemented via software may be written in a high-level procedural language such as object oriented programming or a scripting language. Accordingly, the program code may be written in C, C++, SQL or any other suitable programming language and may comprise modules or classes, as is known to those skilled in object oriented programming. At least some of the elements of the methods described herein that are implemented via software may be written in assembly language, machine language or firmware as needed. In either case, the program code can be stored on a storage media or on a computer readable medium that is readable by a general or special purpose programmable computing device having a processor, an operating system and the associated hardware and software that is necessary to implement the functionality of at least one of the embodiments described herein. The program code, when read by the computing device, configures the computing device to operate in a new, specific and predefined manner in order to perform at least one of the methods described herein.

[00197] Furthermore, at least some of the methods described herein are capable of being distributed in a computer program product comprising a computer readable medium that bears computer usable instructions for one or more processors. The computer readable medium may be provided in various non-transitory forms and other forms such as, but not limited to, one or more diskettes, compact disks, tapes, chips, USB keys, external hard drives, wire-line transmissions, satellite transmissions, internet transmissions or downloads, magnetic and electronic storage media, digital and analog signals, and the like. The computer-useable instructions may also be in various forms, including compiled and non-compiled code.

[00198] The above disclosure generally describes the present application. A more complete understanding can be obtained by reference to the following specific examples. These examples are described solely for the purpose of illustration and are

- 53 -

not intended to limit the scope of the application. Changes in form and substitution of equivalents are contemplated as circumstances might suggest or render expedient. Although specific terms have been employed herein, such terms are intended in a descriptive sense and not for purposes of limitation.

- 5 **[00199]** The following non-limiting examples are illustrative of the present disclosure.

Examples

Example 1

- 10 **[00200]** The media used in these experiments contained 5% FBS; 50% media is media diluted in PBS (1:1).

[00201] RNA was isolated from a core biopsy of a MCF-7 tumour xenograft that has been kept at room temperature in saline for 24 hours where autolytic degradation takes place and separated using an Agilent 2100 Bioanalyzer.

- 15 **[00202]** RNA was isolated from ovarian cancer cell line A2780 cells which had been treated with 10 μ M docetaxel in 50% media or 20 μ M epirubicin, each in 50% media. Cells were treated with the chemotherapeutic for 24 hours at 37°C. Isolated RNA was separated by microcapillary electrophoresis using the Agilent 2100 Bioanalyzer.

- [00203]** **FIGS 1-3** show electropherogram traces of the MCF-7 tumour RNA kept
20 at room temperature with saline resulting autolytic RNA degradation (**FIG. 1**) and A2780 tumour cell RNA treated with chemotherapy drugs docetaxel and epirubicin (**FIG. 2** and **FIG. 3** respectively).

- [00204]** Xenograft tumour kept in saline for 24 hours shows a large peak of small fragments at approximately the 26 second mark of the trace (**FIG. 1**). There is also a
25 loss of area of the 28S peak. Cancer cells treated with chemotherapeutic drugs show a different pattern of RNA disruption. For example, **FIG. 2** which shows human ovarian cancer cell line A2780 treated with chemotherapeutic docetaxel, shows an intermediate peak and a peak adjacent to the 18S peak. **FIG. 3** shows A2780 cells treated with chemotherapeutic epirubicin. The electropherogram shows intermediate bands between

- 54 -

the 18S and 28S peaks. In addition, the location of the detected 28S and 18S peaks are shifted relative to the 18S and 28S peaks compared to intact RNA and autocatalytic RNA degradation.

[00205] RNA isolated from xenograft tumours stored in liquid nitrogen, RNA Later™ solution or left in saline for 1h, 2h, 5h and 24h to induce autocatalytic degradation was analyzed by measuring an RNA Integrity Number (RIN) using Agilent 2100 Bioanalyzer. **FIGS. 8B-8M** shows that the 28S and 18S peaks are not misidentified by this software in RNA samples with autolytic degradation (samples 4-7 and 9-12 of **FIG. 8A**). Circles above the gel lanes indicate that an RIN number could be calculated (**FIG. 8A**). Circles encapsulated by squares indicate samples for which an RIN number could not be calculated. Control samples were untreated xenograft tumours (1-7). Treated samples (8-12) were given docetaxel.

Example 2

[00206] The MA22 RNA electropherogram data was examined to determine which components of the electropherogram best correlated with clinical response. It became evident that the Agilent Bioanalyzer software misidentified either the 28S peak or the 18S peak on electropherograms in >20% of the samples. In these samples, a small portion of the tip of a peak was identified as the whole peak. The Agilent 2100 Bioanalyzer software (Agilent Expert) (Mueller 2004; Vespucci 2005) was able to correctly identify the 28S and 18S peaks on electropherograms of intact RNA which are identifiable by eye. However, in samples with disrupted RNA, electropherogram peaks were misidentified by Agilent RIN software although the correct peaks remained in appropriate locations visually on the electropherogram. The Agilent method which generates RIN values, hereafter referred to as the RIN algorithm, was based on RNA samples which were either intact or had undergone complete autolytic degradation and had few "partially degraded" samples (Schroeder, Mueller et al. 2006).

[00207] Accordingly, methods were used to determine an RDA score herein, as described with respect to **FIGS. 5A and 5B**, which are different from the methods used to calculate RIN. The output RDA scores were subjected to Receiver Operating

Characteristic(ROC) analysis to generate an ROC curve from which the Area under the Curve (AUC) was calculated (Hurley 2011). The ROC curve was then used to determine the utility in discriminating Responders from Non-Responders using certain features of the electropherogram (as measured by a pathological complete response post-treatment).

[00208] *Assessment of RNA quality and concentration using an Agilent 2100 BioAnalyzer with Agilent RNA 6000 Nano kits and Caliper Technology's RNA Nanochips:* The procedure used for assessing the quantity and integrity (also referred to as RNA quality) of the above RNA preparations involved capillary electrophoresis on an Agilent 2100 Bioanalyzer (Agilent Technologies, Mississauga, ON) using a protocol described in detail in the Agilent RNA 6000 Nano Kit Guide available on the Agilent Technologies website (Agilent Publication Part Number: G2938-90035).

[00209] The protocol document includes a detailed description on setting up the assay equipment, preparing and running the RNA Nanochips, and analysis of the capillary electrophoretic data using the Agilent 21000 Bioanalyzer and its associated "Expert software". The document includes "Essential Measurement Practices" to be followed. RNA 6000 "Nanochips" and associated solutions are obtained in RNA 6000 Nano kits that can be purchased from Agilent Technologies (Mississauga, ON). RNA Nanochips are manufactured by Caliper Life Sciences (Hopkinton, MA). The sizes of the rRNAs, and the concentration of RNA in a given sample are determined by extrapolation from a standard curve of reference RNAs provided in the RNA 6000 NANO kits. Data from the Bioanalyzer runs were stored as PDF and XAD files. The Agilent 2100 Expert software was used to obtain all data from the capillary electrophoresis runs, including the size of RNAs and the RNA integrity/quality and quantity for a given RNA preparation. Raw data from each electropherogram was then exported from the XAD files into EXCEL files which contain the raw electropherogram data at time intervals of 0.05s.

Peak Identification Code

[00210] Based on the methods described herein, the current error rate for peak misidentification is 3% for chemotherapy treated samples.

Results of Area and Ratio Calculations

[00211] The error rate for the correct identification of 28S and 18S peaks fell from 25% to 3% using the described method. Included as errors were 28S and 18S peaks that were merged into the “intermediate banding” region resulting in small values for the area under the peak. Including only the mid-identification of the 28S and 18S peaks in electropherograms in which the RNA ran aberrantly gave an error rate of 1%. All ratio values of “0” were eliminated using this method and samples that were assigned RIN values of N/A by the Agilent Bioanalyzer were then quantifiable.

10

TABLE 1: Error Types within MA22 Electropherograms

Type of Error
Peak before 28S
All peaks shifted
Large amount of lower banding with absence of 18S and 28S
Large intermediate peak with absence of 28S and 18S
Large intermediate peak with merged 28S and 18S

[00212] Samples used to assess RDA were taken after the third cycle of chemotherapy. The analysis of pre-therapy samples found that RNA integrity/quality and quantity were generally high and could not discriminate between Responders and non-Responders. *In vitro* studies and treatment of tumours in xenograft mouse models suggest that assessment of RNA disruption could demonstrate an effect as early as after the first cycle of chemotherapy.

15

[00213] The ratio of 28S:18S was calculated based on new area values. In intact RNA the accepted maximum value of 28S:18S is 2.0 (Sambrook, Fritsch, & Maniatis, 1989). This ratio has been shown to change in aging (Mori, Mizuno et al. 1978) (Payao, Smith, Winter, & Bertolucci, 1998), in diseases such as Alzheimer's (da Silva, Payao, Borsatto, Bertolucci, & Smith, 2000) and atherosclerosis (Martinet, De Meyer, Herman, & Kockx, 2004).

20

[00214] Previous work has indicated that during autolytic RNA degradation the height of the 28S peak decreases more rapidly than the height of the 18S peak (Schroeder, et al., 2006). However, in chemotherapy disrupted RNA, this ratio can decrease indicative of a preferential loss of 28S and was also found to increase
5 indicating an increased rate of fragmentation of the 18S peak compared with the 28S peak. This ratio may be useful to discriminate between intact and chemotherapy disrupted RNA but would require two cutoff values such that values greater than approximately 2.5 or less than approximately 1.3 are indicative of chemotherapy disrupted RNA. An increased ratio is not typically found with autolytic degradation.

10 **[00215]** Abnormal 28S:18S ratios have been found in other systems such as connective tissue and tumour tissue (Skrypina et al. 2003) as well as in pure spermatozoal fractions in which 28S depletion results in a significantly lower ratio (median 0.11) than that described for somatic cells (Cappallo-Obermann, Schulze et al. 2011). It has been suggested that this loss in 28S is due to preferential cleavage within
15 GC-rich regions (Johnson, Sandler, Lalancette, Hauser, Diamond, & Krawetz, 2011).

[00216] The area of the intermediate banding region was measured and ratios calculated to normalize to the total area (i.e. intermediate/total area) and to determine how the area of intermediate banding region compares with the area of the 28S and 18S peaks (i.e. intermediate/(28S+18S)). The area in the low banding region (below the 18S
20 peak) was also calculated. Other ratios were also calculated including intermediate/28S, intermediate/18S, (low+intermediate)/(28S+18S) and low banding region/total area. Minimum and maximum ratios were determined for several electropherograms of samples that were taken at the same time and graphed against concentration, an example of which is shown in **FIG. 6A**. Accordingly, if a feature uses the term maximum
25 or minimum then this means that a ratio was calculated for each electropherogram dataset and then the maximum or minimum of these ratios was calculated as the case may be.

[00217] In the ratios of the features that follow below, the term "intermediate" means the area in the intermediate region, the term "28S" means the area of the 28S
30 peak, the term "18S" means the area of the 18S peak, the term "low banding" means the area of the low banding region and the term "total area" means the sum of the areas of

the low banding region, the 18S peak, the intermediate region and the 28S peak. The term concentration (for the graphs in **FIGS. 6A-6D** and **FIGS. 6F-6K**) means the concentration of the RNA in the samples from which the electropherogram was generated. In the cases where the maximum or minimum was taken of a ratio, the concentration used in the plot was that of the electropherogram dataset that had the maximum or minimum value for the ratio.

[00218] The AUC value for the feature of maximum intermediate/(28S+18S) was 0.81 and the AUC value for the ratio of minimum intermediate/(28+18) was 0.75 indicating that both of these ratios are able to discriminate between Responders and Non-Responders. The AUC was the area under the ROC curve. The ROC curve was obtained by applying thresholds to the RDA score that was generated based on the indicated feature. A log-log graph of the first combination of features versus log concentration is shown in **FIG. 6A** and a semi-log graph of the second combination of features versus log concentration is shown in **FIG. 6B**.

[00219] The AUC value for the feature of maximum intermediate/total area was 0.71 and the AUC value for the ratio of minimum intermediate/total area the AUC was 0.56. A semi-log graph of the feature of maximum intermediate/total area versus log concentration is shown in **FIG. 6C**.

[00220] The AUC value for the feature of minimum Low Banding/Total Area was 0.63; however the AUC value for the feature of maximum Low Banding/Total Area value was 0.50. A semi-log graph of the feature of minimum low banding/total area versus log concentration is shown in **FIG. 6D**. The feature of low banding/Total area may be a useful feature in combination with other components as shown in **FIG. 6E**.

[00221] Ratios were also generated for the area of the 28S peak over (i.e. with respect to or divided by) total area (28S/total area) and the area of the 18S peak over total area (18S/total area). The features of maximum and minimum ratios of 28S/total area were graphed against log concentration on semi-log graphs as shown in **FIGS. 6F-6G**, respectively. The features of maximum and minimum ratios of 18S/total area were graphed against log concentration on semi-log graphs as shown in **FIGS. 6H-6I**, respectively.

[00222] The AUC value for the feature of maximum 28S/Total Area was 0.69. The AUC value for the feature of minimum 28S/Total Area was 0.62. The AUC value for the feature of maximum 18S/Total Area was 0.76. The AUC value for the feature of minimum 18S/Total Area was 0.68. This suggests that these features may be useful components to separate Responders from Non-Responders.

[00223] The features of maximum and minimum ratios of Intermediate/28S are shown in **FIGS. 6J-6K**.

[00224] The 18S and 28S peak widths were calculated for the MA22 dataset and assessed for their ability to discriminate between Responders and Non-Responders. The accuracy (e.g. (the number of data whose true labels equal to their predicted labels)/(the total number of data)) as determined by linear discriminant analysis of using this combination of features to generate the RDA score was 0.6 and had an AUC of 0.67.

RNA Disruption Vs RNA Autolytic Degradation

[00225] Distinguishing between the fragmentation effect of drugs and autolytic degradation may be possible using A) changes in electropherogram graphical baseline (i.e. measure of the base between the peaks), B) appearance of low molecular weight banding, C) low RNA concentration and/or D) measuring loss of 28S compared with 18S peaks.

Discussion

[00226] As reported herein the electropherograms of autolytic RNA degradation can differ from chemotherapy induced RNA disruption. Existing tools for analyzing RNA degradation have not been directed to assessing non-autolytic degradation. Accurate analysis of the RNA electropherograms is key to determining the extent of disruption in tumour RNA samples that have been treated with chemotherapy. The disclosed peak identification method provides the ability to fully analyze an increased number of samples since samples that were previously marked an RIN of N/A by the Agilent Bioanalyzer were assessed and area values generated using the methods described

herein. Features associated with rRNA disruption by chemotherapy are listed in **TABLE 2**.

TABLE 2: Features Associated with rRNA Disruption by Chemotherapy

Feature	Effect
28S peak	Loss of area, peak height, peak width
18S peak	Loss of area, peak height, peak width
28S:18S ratio	Increase or decrease
Intermediate banding	Area, peak height, discrete banding?
Low banding	Area, peak height, discrete banding

5

[00227] The maximum ratio of 28S to 18S has been commonly assigned the value of 2.0; and a fall in this ratio has been ascribed to degradation (e.g. less than 1.8). In the study of **Example 2**, it was found that in samples with disrupted RNA, the 28S:18S ratio either decreased, or in some cases increased. In both cases, strong intermediate banding was present. The extent of fragmentation of the 28S differed from the fragmentation occurring in the 18S rRNA.

[00228] The intermediate banding region appears to be an important component in the RNA electropherogram when determining the extent of fragmentation of RNA (e.g. RNA disruption). The appearance of banding in the intermediate region along with 28S and/or 18S bands that make up smaller proportions of the total area indicate a disruption of RNA.

Example 3

[00229] The need for a prognostic biomarker to assess long term breast cancer chemotherapy efficacy is well established. The RNA Disruption Assay (RDA) is a novel prognostic test for women undergoing neoadjuvant chemotherapy that enables assessment of drug efficacy during treatment. RDA was developed from the findings of a clinical trial for women with locally advanced breast cancer (CAN-NCIC-CTG MA-22). Varying doses of docetaxel and epirubicin were given at two weekly (dose dense) or three weekly intervals. Tumours were biopsied in duplicate at 3 time points; pre-therapy,

25

mid-therapy and post-therapy. Trial endpoints included clinical response, pathological complete response (pCR), disease-free survival and overall survival.

Methods:

5 [00230] RDA is based on analysis of RNA electropherograms generated to assess RNA integrity/quality. At mid-therapy, patients with tumours responding to treatment exhibited a dramatic reduction in RNA quality (RNA disruption). A method to provide an RDA score was then developed by combining various features of the electropherogram to discriminate between those patients who subsequently exhibited
10 pCR and those that did not. This method generates an RDA score that indicates the degree of RNA disruption. A high RDA score is associated with subsequent pCR, while a low RDA score indicates that patients are unlikely to receive long term chemotherapy benefit. For this example, the RDA score was based on the maximum value for the ratio of the area of the intermediate region to the sum of the areas of the 28S and 18S peaks
15 (i.e. $\text{intermediate}/(28\text{S}+18\text{S})$).

Results:

[00231] When applied to patients in the MA-22 study (partial patient details available in Parissenti et al. 2010 and PCT/CA2008/001561; 35 additional patients have
20 been analyzed for a total of 85 patients and in some cases 3 samples per subject were analyzed; PPV AND NPV numbers described herein are based on the full data set) after the 3rd or 4th cycle of chemotherapy, the RDA score (e.g. based on the ratio of the Intermediate Area/(28S+18S Areas)) made the following predictions. Of the 85 patients studied, it predicted 32% to be non-responders. This prediction had a negative predictive
25 value of 0.99 with a 95% confidence limit of 0.98-1.0 and a false negative rate of 2%. Of patients achieving pCR, RDA predicted 87% of having an increased chance of responding. This prediction had a positive predictive value of 0.19 reflecting that some patients had a drug effect but did not achieve a pCR. FIG. 7A illustrates the RDA Zones and identified responders and non-responders and FIG. 7B illustrates the ROC curve.
30 The number of responders is also tabulated below in TABLE 3.

- 62 -

[00232] For example, selecting a threshold of 10.2 for the RDA score, the negative prediction value for mid treatment responders is 0.99. This means that a tumour sample having an RDA of >10, for example, can be predicted to be non-responsive to the chemotherapy treatment with greater than 99% confidence.

- 5 **[00233]** The average standard deviation for samples measured in triplicate and using RDI based on the ratio of Intermediate Area/(28S+18S Areas) was 41. This was for mid-therapy samples and is expected as tumours can be highly heterogeneous.

10 **TABLE 3: MA22 Patient Response**

Clinical Response		Complete Pathologic Response pCR		RDA
Complete:	17 patients	6 Patients	7 Patients	Zone 3
				25 Patients No pCR 7 pCR
Partial:	64 Patients	2 Patients	1 Patient	Zone 2
				25 Patients No pCR 1 pCR
No response:	4 Patients			Zone 1
				27 Patients No pCR

Conclusions:

- 15 **[00234]** Use of RDA in clinical practice will enable identification of patients unlikely to respond to breast cancer chemotherapy. Patients with an RDA score that indicates no long term chemotherapy benefit can then avoid the toxic side-effects of chemotherapy and can be switched to an alternate treatment, which may provide a better clinical outcome as well as health care cost savings.

Example 4

[00235] The methods and assays described herein can be used for assessing the benefit of changing therapy for non-responders. Changing chemotherapy protocols midstream on the basis of patient response may result in improved overall survival as well as improved disease free survival rates.

[00236] Predicting ultimate response rapidly, for example, after one or two cycles of a chemotherapy treatment, could provide useful information for maximizing clinical benefit and/or minimizing toxicity.

[00237] To assess this, patients are treated with one or more different chemotherapy drugs. Core biopsies are taken from the patients at one or more times of before, during and/or after treatment. The biopsies are analyzed using an RNA degradation assay. Responders and non-responders predicted based on the RDA assay are identified. Non-responders are randomized into two or more groups. For example, subjects can be maintained on their same treatment or receive adjuvant or other treatment (e.g. e.g. adjuvant, different drug or other therapy such as surgery or radiation). Outcomes for the predicted "non-responder" groups are compared to assess accuracy of prediction and to assess whether altering treatment (e.g. adjuvant, different drug or other therapy such as surgery or radiation) results in improved outcome.

[00238] Additional markers can be analyzed pre, mid and/or post treatment.

Example 5

[00239] Subjects with ER/PR+/HER2- negative and triple negative tumours are assigned a therapy based on, for example, molecular profiling. For example, an arm could receive neo-adjuvant hormonal treatment and a second arm neo-adjuvant chemotherapy A.

[00240] Responders and non-responders are predicted using an RNA disruption assay method. Non-responders can be randomized to receive different or additional treatment. For example, non-responders in the first arm would either continue on the hormonal treatment or receive chemotherapy or hormonal therapy plus an adjuvant

therapy; and non-responders in the second arm would either continue to receive chemotherapy A or receive chemotherapy B and/or surgery. The outcomes such as disease free survival for all patients and also patients switching therapy would be compared to patients continuing therapy.

5

Example 6

[00241] Cells were plated and treated with docetaxel concentrations ranging from 0.001-40 μ M for 24 to 72 hours to assess the effect on cell number, RNA content, and integrity (measured using a Bioanalyzer 2100 from Agilent Technologies, Inc.).

10 **[00242]** A2780 cells were treated for up to 72 hours with docetaxel at 0.005 μ M and 0.2 μ M. Cell counts were measured at each time-point. Cell numbers remained constant for concentrations of docetaxel greater than or equal to 0.005 μ M (**FIG. 9A**). The use of the asterix (*) in **FIGS. 9A** and **9B** refers to whether the differences between the drug-treated samples and the drug-free control are considered statistically significant
15 (in a two-way ANOVA with Bonferroni post test to determine significance).

[00243] A2780 cells treated for up to 72 hours with docetaxel at 0.005 μ M and 0.2 μ M, were analyzed for the amount of RNA per cell. RNA was isolated and electrophoretically separated using an Agilent Bioanalyzer. The amount of RNA per cell was calculated and shown to increase in a time-dependent and dose-dependent
20 manner. The results are shown in **FIG. 9B**.

[00244] A Recovery Assay was next carried out to determine if A2780 cells could recover from treatment with 0.2 μ M docetaxel. A2780 cells were treated with 0.2 μ M docetaxel for up to 72 hours and then washed and replated in drug-free media. No resumption of cell proliferation was seen under these conditions (**FIG. 9C**).

25 **[00245]** Electrophoretic separation of A2780 cell RNA from cells treated with docetaxel for 24 hours at concentrations ranging from 0 to 40 μ M, demonstrate the appearance of bands in the intermediate region and the low banding region. RNA was isolated and electrophoretically separated on an Agilent Bioanalyzer. As shown in **FIG. 9D** bands appeared just below the 28S (e.g. intermediate) and 18S bands (e.g. low
30 banding region) at 0.005 μ M and peaked at 0.2 μ M.

[00246] Traces of the electropherograms also illustrate that the banding which appears just below the 28S and 18S peaks at a concentration of 0.005 μ M and peaks at a concentration of 0.2 μ M (**FIG. 9E**).

[00247] It should be noted that the RDI values shown in **FIGS. 9F** and **9G** were
5 determined using the methods 200 and 300 although method 400 could be used instead of method 300.

[00248] Analysis using RDA was performed. **FIG. 9F** is a plot of the measure of Intermediate $I/(28S+18S)$ for A2780 cells treated with docetaxel concentrations of 0.005 μ M and 0.2 μ M for 24 to 72 hours. Each value represents a single sample. Two
10 replicate experiments had similar results. The graph demonstrates that the ratio of the Intermediate $I/(28S+18S)$ increases dramatically with time. Similarly, a graph of the ratio of Low banding area/ $(28S+18S)$ demonstrates this ratio also increases with time (**FIG. 9G**). As before each value represents a single sample. Two replicate experiments had similar results.

[00249] A time course experiment was conducted in A2780 cells. Cells were
15 treated for up to 72 hours with docetaxel at 0.005 μ M and 0.2 μ M. RNA was isolated and run on an Agilent Bioanalyzer. **FIG. 9H** demonstrates that banding appears below the 28S and 18S bands at all time-points.

[00250] A docetaxel A2780 resistant cell line was made as described in Li et al.,
20 Int. J. Mol. Med 2004 **14**(20): 257-264.

[00251] A time course experiment was conducted. Sensitive parental and docetaxel resistant A2780 cells were treated with docetaxel for 24 hours. RNA was isolated and run on an Agilent Bioanalyzer.

[00252] As shown in **FIG. 9I**, banding appears below the 28S and 18S bands at
25 0.005 μ M in sensitive cells. However, no banding is evident in resistant cells.

[00253] The effect of a caspase inhibitor on RNA integrity changes was next assessed. Cells were treated with or without docetaxel at a concentration of 0.2 μ M in the presence or absence of the caspase inhibitor Q-DEVD-OPH (Biovision Laboratories) at a concentration of 10 μ M.

[00254] RNA was isolated and run on an Agilent Bioanalyzer and the banding pattern is shown in **FIG. 9J**. The inhibitor reduces the loss of 28S and 18S bands at 72 hr. Bands below the 28S and 18S bands are detectable in the presence and absence of the caspase inhibitor.

- 5 **[00255]** Although cell number remained constant for concentrations $\geq 0.005\mu\text{M}$ docetaxel, withdrawal of the drug did not result in resumption of cell proliferation. RNA content increased per cell ($p < 0.05$) while RIN did not change significantly in this range. However, novel discrete bands appeared in the rRNA banding pattern at $0.005\mu\text{M}$, peaking at $0.2\mu\text{M}$ docetaxel, just below the 28s and 18s rRNA bands. In contrast,
- 10 docetaxel-resistant A2780DXL cells did not display similar changes upon treatment, indicating that changes in tumor cell RNA content and integrity could be used to monitor response to chemotherapy agents.

Example 7

- 15 **[00256]** A2780 ovarian cancer cells were treated with radiation of 2 to 10 Gray using a Gulmay RS320 Irradiation System and subsequently harvested after 24, 48 and 72 hr. RNA was isolated and run on an Agilent Bioanalyzer. Analysis of the electropherogram demonstrated that an increase in the ratio Intermediate Area/(28S+18S Areas) at 72 hours is detectable in radiation treated cells (**FIG. 10A**). For
- 20 example, cells treated at 10 Gray for 72 hours show RNA disruption in the intermediate region and the low banding region, but not much difference in the area of the electropherogram where the autolysis peak resides. **FIG. 10B** shows A2780 cells treated with 5 Gy for 24 hrs (lane A), 5 Gy for 48 hrs (lane B), 5 Gy for 72 hrs (lane C), 10 Gy for 24 hrs (lane D), 10Gy for 48 hrs (lane E) and 10 Gy for 72 hrs (lane F). RNA
- 25 disruption is evident at 72 hrs at both dose levels.

Example 8**Methods****Administration of the FEC-D regimen with concurrent radiation**

[00257] Thirty two patients with stage III non-metastatic, non-inflammatory locally advanced breast cancer were treated with neoadjuvant 5-Fluoro-uracil, Epirubicin, and Cyclophosphamide (FEC also referred to as CEF) q3 weekly for 4 cycles followed by weekly Docetaxel (35 mg/m²) concurrently with regional radiation (45 Gy with 16 Gy boost in 25 & 5 fractions) for 6 weeks followed by an additional 3 weeks of docetaxel chemotherapy without radiation. This was followed by a modified radical mastectomy. Patient and tumour characteristics were recorded at baseline and following treatment and clinical response and treatment-related toxicities noted. Image guided serial 14 gauge tumour core biopsies were taken from the patients pre-, mid- and post-treatment, and 1 mm³ sections were immediately taken from the biopsies, immersed in RNAlater™, and stored frozen. MID treatment is after FEC but before docetaxel with concurrent radiation therapy.

Isolation of RNA from Tumour Core Biopsies

[00258] RNA was isolated from image-guided tumour core biopsies of patients pre-, mid-, and post-treatment using Qiagen miRNeasy® Mini kits, following a modification of the protocol published on the manufacturer's website, <http://www1.qiagen.com/literature/handbooks/literature.aspx?id=1000291>. Biopsies were cut into several sections for various assays, with the section used for RNA integrity analysis placed in RNAlater. The biopsies in RNAlater were immediately dropped in 0.5 ml of RLT buffer containing β-ME (10 µl into 1 ml) in a 1.5 ml tube. The biopsies in RLT buffer were homogenized with a Coreless motor homogenizer for 5 min (from the Kontes Glass Company). The lysate was then passaged at least 5 times through a 20-gauge needle (0.9 mm diameter) fitted to an RNase-free syringe. The sample was then centrifuged at high speed in a refrigerated microfuge at 4°C for 3 minutes, with transfer of the supernatant to a new tube. One volume (500 µl) of 70% ethanol was then added to the supernatant and the sample was mixed well by repeated pipetting. If some lysate

- 68 -

- was lost during homogenization, then the volume of ethanol was adjusted accordingly. Visible precipitates formed after the addition of ethanol in some samples did not affect the RNA isolation procedure. A maximum of 700 µl of the sample, including any precipitate, was added to a Qiagen® mini column and placed in a 2 ml collection tube.
- 5 The column was centrifuged for 15 s at $\geq 8000 \times g$ ($\geq 10,000$ rpm) and the flow-through discarded. The remainder of the sample was then added to the column and the column centrifuged again. From this point forward, the column was then washed twice in RPE buffer and dried by centrifugation as per the manufacturer's protocol. The RNA was then eluted from the column in 30 µl of RNase-free water and the elute reapplied and
- 10 eluted from the column to increase the yield and concentration of the RNA obtained.

Assessment of RNA quality using an Agilent 2100 Bioanalyzer

- [00259] The above RNA samples were applied to RNA 6000 Nano Lapchips™ (purchased from Agilent Biotechnologies, Inc.) and subjected to capillary electrophoresis
- 15 using an Agilent® 2100 Bioanalyzer.

[00260] Further details of the study and clinical results are provided Cancer Res 2012;72(24 Suppl):Abstract nr P1-14-13.

Results

- 20 [00261] RNA isolated from patient tumour samples after radiation was run on an Agilent Bioanalyzer. Analysis of the electropherogram demonstrated an increase in the ratio Intermediate Area/(28S+18S Areas) in two patients that had a pathological complete response (FIG. 11a).
- [00262] Analysis of the electropherogram also demonstrated an increase in the
- 25 ratio Low Area/(28S+18S Areas) in two patients that had a pathological complete response (FIG. 11b).

Example 9

[00263] Another method can be used for peak identification and generally comprises determining a standard sample (i.e. a normal sample) for a plurality of samples, for example such as the samples provided on a platform such as an RNA chip and the 28S and 18S peaks, as well as other features of interest, are identified by comparing the retention times for each sample to the standard sample. For example, when the standard sample is determined the method then generally comprises determining which samples require adjustment (e.g. forming an adjustment group), determining the standard retention time for the platform, adjusting samples in the adjusting group based on the standard retention time, determining peaks using the standard retention time and calculating one or more features of the peaks and nearby regions (as shown by the examples given in the discussion of **FIG. 5A**). By identifying a "standard" or "normal" sample for each platform, (for example, when multiple test samples are run simultaneously, as in the case of when an Agilent Bioanalyser chip platform can run 12 samples per chip, for example), the 28S and 18S peaks can more readily be determined by comparing the retention times of each sample to those of the "standard" sample of the chip platform. This then allows for a normalization for the samples that are collected at the same time using the same means. Further details on this method are provided in **Example 10** and **FIG. 12A** which provides a flowchart of an example embodiment to implement the aforementioned method of peak identification and feature calculation.

[00264] Additional features have also been identified as can be seen in **FIG. 12B** in which the low banding region is trisected into several regions including the low A banding region, the low B banding region and the low C banding region.

Example 10Modified Method using the Agilent Bioanalyzer

[00265] Referring now to **FIG. 12A**, shown therein is an example of an alternative embodiment of a method 400 for determining peaks and calculating features for electropherogram datasets. The method 400 generally involves identifying a standard sample or normal sample for a plurality of electropherogram datasets obtained from

- 70 -

samples on a common analysis platform, such as an RNA chip for example, and then using the standard sample to adjust the location of peaks for the electropherograms for the samples that are determined to require adjustment and are therefore contained within an adjustment group. The plurality of electropherogram datasets are generated using samples that are analyzed using the same platform such as an RNA chip, for example. It has been found that the peak retention times are different for different RNA chips (e.g. different platforms). It has been found that the method 400 allows for the more accurate assessment of the 18S peak and the 28S peak in samples that are highly fragmented.

- 10 **[00266]** At 402, the method 400 determines a range for the peaks of interest. This range may be initialized to a default setting and then shifted if required. For example, for the 18S and 28S peaks, the default ranges can be initially set to [39.5 s, 44.95 s] and [45.05 s, 53.5 s]. On a scale of 0 to 100, the range for the 18S region is [45.85, 61.25], and that for the 28S region is [61.6, 85.75]. These ranges can then be shifted under
 15 certain conditions, such as, but not limited to the case when the marker region is not at its expected location. For example, these ranges may be shifted if the time of the marker, which is the time associated with the first peak that is a dye-only peak meaning that it does not contain any RNA but rather indicates the start of the run for the gel, is not at about 22.5 seconds. For these samples, a shift factor is calculated according to
 20 equation 1:

$$rngshft = marker\ time - 22.5 \quad (1)$$

- where marker time is the start time of the marker. The range for the 18S region is then shifted up according to $[39.5 + rngshft \times 1.5, 44.95 + rngshft \times 1.5]$ and the range of the for the 28S region is shifted up according to $[45.05 + rngshft \times 2.5, 53.5 +$
 25 $rngshft \times 2.5]$.

[00267] Referring now to FIG. 12C, shown therein is a graph of an example peak which is now used to illustrate some aspects of how peaks are defined using method 400. FIG. 12C shows a plot of a general peak with one sub-peak with the x-axis being time in seconds and the y axis being is the value [FU].

- 71 -

- [00268]** It should be noted that a real-valued function f defined as a non-linear line is said to have a local (i.e. a relative) maximum point at the time point x^* , if there exists some $\varepsilon > 0$ such that $f(x^*) \geq f(x)$ when $|x - x^*| < \varepsilon$. The value of the local maximum point does not have to be larger than the value of all the other points but rather it just needs to be the largest value within a local range. The function f has a global (i.e. absolute) maximum point at x^* , if $f(x^*) \geq f(x)$ for all x . A global maximum point must have the largest value. A function can have more than one local maximum point, but only one global maximum point. In a similar way, there is also a local minimum point and a global minimum point.
- 10 **[00269]** The whole peak region is defined from time **a** to time **b**. The time point **a** is the starting point of the peak and the time point **b** is the ending point. The global minimum point in the range **[a, b]** is either **a** or **b**. The peak point is defined as the global maximum point, i.e. **c**. The peak height is defined as the value of the highest point, i.e. the value of **c**. The peak width is defined as (*ending time – starting time*), i.e. **b-a**.
- 15 The peak area is defined as the area above the x-axis and under the curve from time point $x = a$ to $x = b$. The local maximum point **e** is a sub-peak point. The 18S and 28S peaks can have more than one sub-peak. The local minimum point **d** is the sub-peak low.
- [00270]** At 404, the peaks are located by first searching for all possible peak candidates in the two identifying ranges. This involves searching for local maximums surrounded by local minima on either side. Each part of the electropherogram within the range which has these criteria is designated as a “peak candidate”. The *height difference* of each possible peak candidate is then calculated. The *height difference* of a peak candidate is defined as the difference of the highest peak value and the average lowest peak values. The lowest peak values that are averaged are the starting point and ending point of the peak at the starting time and the ending time for each peak and any sub-peak low values (i.e. local minima) are not included. The two peak candidates with the largest *height differences* can be chosen as the 18S and 28S. The 18S peak or the 28S peak may have the higher value. Once the 18S peak and the 28S peak are correctly identified, the desired features can then be calculated such as, but not limited to, the width, the height and the area of the 18S and 28S peaks, the intermediate area, the low
- 20
- 25
- 30

C banding region area, etc. The sub peak is included in the area calculation or the width calculation and it has no impact on the height (because it is always smaller). The values for these features can be calculated now since they will not change for some of the samples. In an alternative embodiment, the values for these features may be calculated
 5 for all samples after the samples requiring adjusting have been adjusted.

[00271] It should be noted that with method 400, there are two cases when the peak is allowed to have more than one maximum value. First, if the width of a peak is less than 0.15 seconds, this peak is considered as a sub-peak of its adjacent peak. Secondly, if the difference between the starting value and the ending value of a peak is
 10 too large, this peak is considered as a sub-peak of its adjacent peak. The difference is considered to be too large when the height of the sub-peak (i.e. the value at time point e minus the value at time point b in **FIG. 12C**) divided by the difference between the local maximum and local minimum (i.e. the value at time point e minus the value at time point d in **FIG. 12C**) is larger than a threshold value, such as, but not limited to, 5, for
 15 example, then it is too large. In general it has been found that one or two peaks may be associated with the 18S peak and the 28S peak using method 400.

[00272] Furthermore, in method 400, there is no restriction based on the width of the 18S peak and the 28S peak when determining the 18S peak and the 28S peak locations. Also, in method 400, there is no restriction based on the distance between the
 20 18S peak and the 28S peak when determining the 18S peak and the 28S peak locations.

[00273] At 406, the standard sample for the plurality of samples is determined. The standard sample may be defined to be the sample having the smallest *standard score* in the plurality of samples wherein the *standard score* is defined equation 2.

$$25 \quad \text{standard score} = \frac{\text{intermediate area} + \text{lower b area} + \text{lower c area}}{\text{18S peak area} + \text{28S peak area}} \quad (2)$$

[00274] At 408, the retention time of the standard sample is compared with the retention time of the other samples. The retention time is examined for both the 28S and 18S peaks. Those samples whose retention time is not around the retention time of the standard sample will be assigned to an adjustment group, and the rest of the samples
 30 will be assigned to a standard group. A threshold value can be used to assess this (i.e.

- 73 -

the retention time not "being around") such as, but not limited to, 0.5 seconds. Accordingly, if the difference between the retention time of both the 18S peak and the 28S peak for a sample is off by more than 0.5 seconds compared to the same peaks of the standard sample, then the sample is put into the adjustment group and requires
5 adjusting.

[00275] At 410, the standard retention time for the 18S peak and the 28S peak is chosen based on the samples in the standard group. For example, the standard retention time may be chosen as the median of the retention time of the samples in the standard group. In alternative embodiments, other statistical measures may be used
10 such as, but not limited to, the mean or some variations on the mean.

[00276] At 412, the 18S and 28S peaks for each sample in the adjustment group are determined by locating the peaks that are around the standard retention times determined in act 410 (i.e. the times of the 18S and 28S peaks determined from the standard samples). For example, using the samples that are in the standard group, and
15 hence represent the most intact RNA, the median value of the retention time for the 28S and the 18S is identified. Each sample in the adjustment group is then examined to find the peaks that appear at that retention time. However, the actual retention time of the 28S and 18S peaks may not be identical to the those of the standard sample as the peak may not have its highest point at the same exact retention time. In this case if the
20 peak area still includes the "standard retention time", this peak is designated as the 18S peak or the 28S peak even though the maximum value of the peak does not occur exactly at the same retention time. All measurements (width, height, area) are then taken based on this peak. Other regions of interest can also be determined once the 18S and 28S peaks are located such as, but not limited to, the low C banding region and the
25 intermediate region, for example.

[00277] At 414, the features of interest that are used in determining the RDA score or RDI are then calculated for all of the samples.

[00278] An RDI score may be calculated for each sample on the RNA chip including the "control" sample, which may be defined as the sample on the RNA chip
30 that is the closest to "intact RNA".

- 74 -

[00279] In one embodiment, the combination of features is a ratio of the low C banding region area to the sum of the 18S peak area and the 28S peak area, which can be represented by the expression "low C/(18S+28S)". Method 300 can also be used with this combination of features.

- 5 **[00280]** It should be noted that the method 400 can also operate if there are two samples on the RNA chip in which one of the samples is a positive control sample of intact RNA, which in fact is run on every chip, and the other sample can be the sample for which an RDA or an RDI value is obtained. Alternatively, the second sample may also be an analyzing sample which may be useful under certain circumstances.
- 10 **[00281]** Accordingly, at least one embodiment described herein comprises obtaining two electropherogram datasets corresponding to a test sample and a control sample, the test sample being a unique biological sample comprising cellular RNA optionally at a time point before, during or after the treatment requiring analysis and the control sample being another unique biological sample or a positive control sample of
- 15 intact RNA or approximately intact RNA; placing the electropherogram of the test sample into an adjustment group and placing the electropherogram of the control sample into a standard group; identifying normal characteristics representative of a normal sample from the standard group; determining locations of peaks for the electropherogram in the adjustment group and adjusting the location of the peaks based on the normal
- 20 characteristics; and determining values for features from the adjusted electropherogram and optionally the electropherogram from the standard group when the control sample is another unique biological sample requiring analysis.

Example 11

- 25 *The Boundaries of the Low A, Low B, and Low C banding regions*

[00282] The boundaries were set based on the location of discrete sets of peaks within the Low region, which is the region between the marker region and the 18S region (i.e. 18S band). The Low A banding region may be defined as the range between the marker region and 30.0s. The Low B banding region may be defined as the range

- 75 -

between 30.0s and 35.0s. The Low C banding region may be defined as the range between 35.0s and the 18S region.

[00283] Linear Discriminant Analysis was used to identify features that are most predictive in the MA22 data set. Combinations of two or three features were found to improve prediction performance and many combinations listed in **TABLES 7 to 8** were very close in their ability to predict responders. All peak measurements, area measurements, and width measurements for various features (i.e. 28S, 18S, intermediate, Low C) used in this analysis were normalized to the entire area. The entire area is defined to be the sum of the areas of the low banding region, the 18S region, the intermediate region and the 28S region which represents the total amount of RNA present in the analyzed sample.

Example 12

[00284] The methods 200 and 400 of were used to assess the MA22 trial dataset in a similar fashion as for the use of methods 200 and 300 in **Example 2**. RNA isolated from mid-therapy tumour samples was analyzed using an Agilent Bioanalyzer. The result was that the misidentification of the 28S peak in the chemotherapy treated samples fell to 0.3% using the method 400 (see **TABLE 4**).

TABLE 4: Error Rates and Types in MA22 Dataset

Reason	Rate
Inadequate Sample	9.6% of samples
Aberrant Run - baseline or marker issue	1.9% of samples
Mis-identification of 28S and 18S peak	0.3% of samples

[00285] The ratio of the intermediate area to the sum of the 28S and 18S areas was generated for mid-therapy tumour samples and plotted versus concentration on a log-log graph. In particular, the maximum ratio value for each patient was graphed (**FIG. 13**). RDA zones were established based on clinical criteria. In this example, the clinical

- 76 -

criteria was a negative predictive value for the threshold value between zone 1 and 2 at 0.98 or 0.99. For zones 2 and 3, the clinical criteria were chosen to be the best PPV value possible while including most of the responders in zone 3. The maximum RDA for each patient was graphed vs. log concentration on a semi-log chart and is shown in **FIG.**

5 **14A** for the various RDA zones. **FIG. 14B** shows the Receiver Operating Characteristic (Sensitivity vs. 1-Specificity) curve for Maximum value of the ratio of the Intermediate area/(18S+2S). The Area under the curve (AUC) was about 0.8.

[00286] A ratio of the Low C banding region area to the sum of the 28S and 18S areas was also generated. The maximum value for this ratio for each patient was
10 plotted vs. RNA sample concentration (for the sample that had the maximum ratio) on a log-log plot and is shown in **FIG. 14C**.

[00287] A ratio of the sum of the intermediate area and the Low C banding region area to the sum of the 28S and 18S areas was also generated. The maximum value for this ratio for each patient was plotted vs. RNA sample concentration (for the sample that
15 had the maximum ratio) and is shown in **FIG. 14D**.

[00288] Other ratios were also calculated and then Linear Discriminant Analysis (LDA) was used to identify the features that were most predictive for the MA22 data set. The details for the application of the LDA are discussed in **Example 13** below. In general, combinations of two or three features were found to provide superior prediction
20 performance but most of the combinations listed in **TABLES 8** and **9** below were found to be close in their ability to predict responders vs. non-responders.

[00289] **TABLE 5** identifies features associated with RNA disruption by Chemotherapy determined while using method 400.

- 77 -

TABLE 5: Features Associated with rRNA Disruption by Chemotherapy identified using Linear Discriminant analysis

Feature	Effect
28S peak	Loss of area, peak height, peak width
18S peak	Loss of area, peak height, peak width
Intermediate region	Area, discrete banding
Low C region	Area, discrete banding
Concentration	Decrease
Ratio of 28S to 18S	Increase or decrease

- 5 **[00290]** When the methods 100 and 400 were applied to patients in the MA-22 study (patient details available in Parissenti et al. 2010 and PCT patent application No. PCT/CA2008/001561) after the 3rd or 4th cycle of chemotherapy, the RDA score was used to predict that, out of the 85 patients that were studied, 27% were non-responders. This prediction had a negative predictive value of 0.99 with a 95% confidence limit of
- 10 0.95-1.0. Of patients achieving pCR, the RDA score predicted 87% of these patients as having an increased chance of responding. This prediction had a positive predictive value of 0.22 reflecting that some patients had a drug effect but did not achieve a pCR. **FIG. 14A** illustrates the RDA Zones and identified responders and non-responders and **FIG. 14B** illustrates the ROC curve. The number of responders is also tabulated below in
- 15 **TABLE 6.**

TABLE 6: MA22 Patient Response

Clinical Response		Complete Pathologic Response pCR		RDA
Complete	17 patients	6 patients	7 patients	Zone 3
				38 patients 31 No pCR 7 pCR
Partial	64 patients	2 patients	1 patient	Zone 2
				24 patients 23 No pCR 1 pCR
No response	4 patients			Zone 1
				23 patients No pCR

Example 13

- 5 **[00291]** The Area Under Curve (AUC) is an important measure of accuracy for a detection or prediction technique as is known to those skilled in the art. Accordingly, the AUC was used to determine which features and which combinations of features are useful for predicting responders or non-responders to a particular treatment based on
- 10 samples taken at a certain time point in the treatment regimen. For example, such samples include patient samples that exhibit RNA changes in response to an inducer such as chemotherapy or radiation therapy. Linear Discriminant Analysis was also used (one of the most commonly used classification methods) with different thresholds to assess different features to determine which features or combination of features had the best predictive or discriminative ability. LDA was used when two or three sets of
- 15 combinations were plotted against one another on two dimensional or three dimensional plots respectively as is shown in **TABLES 8** and **9** and this is referred to as two dimensional feature sets and three dimensional feature sets respectively.

- [00292]** **TABLE 7** provides the AUC values when a one dimensional feature set comprising single features was assessed on the MA22 dataset using the methods 100
- 20 and 400. In this case, the values for the various one dimensional single features were

calculated and compared to various threshold values to discriminate between responders and non-responders (i.e. the threshold values were scalar values).

TABLE 7: AUC for one dimensional feature set analysis on the MA22 dataset

Features	AUC
18S peak area	0.7706
28S peak area	0.7500
18S peak height	0.7184
28S peak height	0.7848
18S peak width	0.4509
28S peak width	0.6266
Intermediate region area	0.7041
Low C banding region area	0.7041
Low B banding region area	0.6978
Concentration	0.5316

5

[00293] When a one dimensional feature set, in which only a single feature is used, is used to discriminate between responders and non-responders, the 18S peak area, the 28S peak area and the 28S peak height were found to have the best performance; while the features of 18S peak height, intermediate area, low C area and low B area had the second best performance. However, the features of 18S peak width, 28S peak width and concentration may also carry some useful prediction information.

[00294] Accordingly, in one embodiment, the single feature may be the 18S peak area in the 18S region. This can be represented by the expressions "18S" or "18S peak area".

15 **[00295]** In one embodiment, the single feature may be the 28S peak area in the 28S region. This can be represented by the expressions "28S" or "28S peak area".

[00296] In one embodiment, the single feature may be the 18S peak height in the 18S region. This can be represented by the expression "18S peak height".

20 **[00297]** In one embodiment, the single feature may be the 28S peak height in the 28S region. This can be represented by the expression "28S peak height".

[00298] In one embodiment, the single feature may be intermediate area. This can be represented by the expressions "Intermediate" or "Intermediate area".

[00299] In one embodiment, the single feature may be low C banding region area. This can be represented by the expression "low C area" or "low C".

5 **[00300]** In one embodiment, the single feature may be low B banding region area. This can be represented by the expression "low B area" or "low B".

[00301] **TABLE 8** provides the AUC values when two dimensional feature sets, comprising at least one feature in each feature set, was assessed on the MA22 dataset
10 using the methods 100 and 400. In this case, the values for the various two dimensional feature sets were calculated and compared to various two dimensional threshold zones to discriminate between responders and non-responders (i.e. the threshold zones are defined by lines).

15 **TABLE 8: AUC for two feature analysis on the MA22 dataset**

Combination of Features	AUC
18S peak area vs. 28S peak area	0.7722
18S peak height vs. 28S peak height	0.7769
(18S peak area+28S peak area) vs. intermediate area	0.7587
(18S peak area+28S peak area) vs. low B banding region area	0.7745
(18S peak area+28S peak area) vs. low C banding region area	0.7951
(18S peak area+28S peak area) vs. (18S peak width+28S peak width)	0.7935
(18S peak area+28S peak area) vs. concentration	0.7832
(18S peak height+28S peak height) vs. intermediate area	0.7508
(18S peak height+28S peak height) vs. low C banding region area	0.784
(18S peak height+28S peak height) vs. (18S peak area+28S peak area)	0.7927
(18S peak height+28S peak height) vs. (18S peak width+28S peak width)	0.7698
(18S peak height+28S peak height) vs. concentration	0.7714

[00302] When a two dimensional feature set was used to do the discrimination between responders and non-responders, it was found that the peak areas and peak heights are useful for discrimination. Accordingly, better discrimination results can be
20 obtained when combining them with other features. From **TABLE 8**, the two dimensional

feature sets of (18S peak area+28S peak area) and low C banding region area, (18S peak area+28S peak area) and (18S peak width+28S peak width), and (18S peak height+28S peak height) and (18S peak area+28S peak area) were found to have the best AUC results. Other combinations listed in **Table 8** may also be useful.

5 **[00303]** Accordingly, in one embodiment, the two dimensional feature sets are used in a two dimensional plot using two orthogonal variables representing each feature set to graph an RDA coordinate that is used to discriminate responders from non-responders. One of the variables is defined by a first feature set and the other variable is defined by a second feature set. These variables may be combined using LDA or QDA
10 as will be described with respect to **FIGS. 15A-15D**. The first variable comprises the 18S peak area and the second variable comprises the 28S peak area. This can be represented by the expression "18S area vs. 28S area" or "18S vs. 28S".

[00304] In another embodiment, the first variable comprises the 18S peak height and the second variable comprises the 28S peak height. This can be represented by the
15 expression "18S peak height vs. 28S peak height".

[00305] In another alternative embodiment, the first variable comprises the sum of the 28S peak area and the 18S peak area and the second variable comprises the low C banding region area. This can be represented by the expression "(28S+18S) vs. low C".

[00306] In another alternative embodiment, the first variable comprises the sum of
20 the 28S peak area and the 18S peak area and the second variable comprises the low B banding region area. This can be represented by the expression "(28S+18S) vs. low B".

[00307] In another alternative embodiment, the first variable comprises the sum of the 18S peak area and the 28S peak area and the second variable comprises the sum of the 18S and 28S peak widths. This can be represented by the expression "(18S+28S)
25 vs. (18S peak width +28S peak width)".

[00308] In another alternative embodiment, the first variable comprises the sum of the 28S and 18S peak heights and the second variable comprises the sum of the 28S peak area and the 18S peak area. This can be represented by the expression "(28S peak height+18S peak height) vs. (28S +18S)".

- [00309]** In another alternative embodiment, first variable comprises the sum of the 18S peak area and the 28S peak area and the second variable comprises the intermediate area. This can be represented by the expression "(18S+28S) vs. (intermediate)".
- 5 **[00310]** In another embodiment, the first variable comprises the sum of the 18S peak area and the 28S peak area and the second variable comprises the RNA concentration of the sample. This can be represented by the expression "(18S+28S) vs. (concentration)".
- 10 **[00311]** In another embodiment, the first variable comprises the sum of the 18S peak height and the 28S peak height and the second variable comprises the RNA concentration of the sample. This can be represented by the expression "(18S peak height + 28S peak height) vs. (concentration)".
- 15 **[00312]** It should be noted that concentration may be used as a feature in two-feature analysis as it has been found to show good discriminative ability when combined with some other features. The concentration is the RNA concentration in the sample which may be determined using various techniques. For example, in embodiments where the electropherogram is produced by a system like the Bioagilent system, the RNA concentration can be determined using measurements obtainable from the electropherogram. In an example, RNA concentration may be determined by comparing
- 20 to a known quantity of a known standard such as the marker or ladder. For example the amount of the RNA in a standard or ladder band is divided by the area under the curve for the entire marker. The Total Area of each sample may then be multiplied by this value to give a RNA concentration for each sample.
- 25 **[00313]** In another embodiment, the first variable comprises the sum of the 18S and 28S peak heights and the second variable comprises the intermediate area. This can be represented by the expression "(18S peak height+28S peak height) vs. (intermediate)".
- [00314]** In another embodiment, the first variable comprises the sum of the 18S and 28S peak heights and the second variable comprises the low C banding region

area. This can be represented by the expression "(18S peak height+28S peak height) vs. low C".

[00315] In another embodiment, the first variable comprises the sum of the 18S and 28S peak widths and the second variable comprises the sum of 18S peak area and the 28S peak area, which can be represented by the expression "(18S peak width+28S peak width) vs. (18S+28S)".

[00316] **TABLE 9** provides the AUC values when three dimensional features sets, comprising at least one feature in each feature set, was assessed on the MA22 dataset using the methods 100 and 400. In this case, the values for the various three dimensional feature sets were calculated and compared to various three dimensional threshold zones to discriminate between responders and non-responders (i.e. the threshold zones were defined by planes).

TABLE 9: AUC for three feature analysis on the MA22 dataset

Combinations of Features	AUC
(18S peak area+28S peak area) vs. (18S peak height+28S peak height) vs. low C banding region area	0.7864
(18S peak area+28S peak area) vs. (18S peak height+28S peak height) vs. (18S peak width+28S peak width)	0.7761
(18S peak area+28S peak area) vs. (18S peak height+28S peak height) vs. intermediate area	0.7445
(18S peak area+28S peak area) vs. (18S peak width+28S peak width) vs. low C banding region area	0.8141

15

[00317] The three dimensional feature sets were selected for analysis based on the best two dimensional feature set combinations. The three dimensional feature set where the first feature comprises (18S peak area+28S peak area), the second feature set comprises (18S peak width+28S peak width) and the third feature set comprises lower c area was found to have the highest AUC based on the sample study. It has also been found that using three dimensional feature sets increases the spread in the data which makes it easier to discriminate between responders and non-responders.

20

[00318] Accordingly, in one embodiment, the three dimensional feature sets are used in a three dimensional plot using three orthogonal variables representing each

- 84 -

feature set to graph an RDA coordinate that is used to discriminate responders from non-responders. One of the variables is defined by a first feature set, another variable is defined by a second feature set and the third variable is defined by a third feature set. These variables may be combined using LDA or QDA as will be described with respect to **FIGS. 15A-15D**. The first variable comprises the sum of the 18S peak area and the 28S peak area and is associated with a first axis. The second variable comprises the sum of the 18S peak height and the 28S peak height and is associated with a second axis. The third variable comprises the low C banding area and is associated with a third axis. This can be represented by the expression "(18S+28S) vs. (18S peak height + 28S peak height) vs. low C banding area".

[00319] In another embodiment, the first variable comprises the sum of the 18S peak area and the 28S peak area and is associated with a first axis. The second variable comprises the sum of the 18S peak height and the 28S peak height and is associated with a second axis. The third variable comprises the sum of the 18S peak width and the 28S peak width and is associated with a third axis. This can be represented by the expression "(18S+28S) vs. (18S peak height + 28S peak height) vs. (18S peak width + 28S peak width)".

[00320] In another embodiment, the first variable comprises the sum of the 18S peak area and the 28S peak area and is associated with a first axis. The second variable comprises the sum of the 18S peak height and the 28S peak height and is associated with a second axis. The third variable comprises the intermediate area and is associated with a third axis. This can be represented by the expression "(18S+28S) vs. (18S peak height + 28S peak height) vs. intermediate area".

[00321] In another embodiment, the first variable comprises the sum of the 18S peak area and the 28S peak area and is associated with a first axis. The second variable comprises the sum of the 18S peak width and the 28S peak width and is associated with a second axis. The third variable comprises the low C band area and is associated with a third axis. This can be represented by the expression "(18S+28S) vs. (18S peak width + 28S peak width) vs. low C band area".

30

Example 14

[00322] The following example describes how a selection threshold is determined for a dataset for a selected set of features comprising the combination of features (18S peak area+28S peak area) and (18S peak width+28S peak width), which is showed
 5 herein for illustrative purposes only. This procedure can be repeated with other features or combinations of features. It should be noted that the feature sets that are used with LDA or QDA have been normalized with respect to the total area which is the area of the low banding region, the 18S region, the intermediate region and the 28S region.

[00323] Referring now to **FIG. 15A**, a first variable representing the 18S peak area+28S peak area was plotted along the x-axis and a second variable representing the 18S peak width+28S peak width was plotted along the y-axis for different patients from the MA22 study. The circles 'o' represent responders and the crosses 'x' represent non-responders in **FIG. 15A**.
 10

[00324] For each threshold value, LDA, which is a classification method, was used to generate values for coefficients that can be used with the first and second variables and the threshold value to define a line that can be used to divide the samples into two groups (i.e. responders and non-responders). This line satisfies equation 3:
 15

$$a \times (18S \text{ peak area} + 28S \text{ peak area}) + b \times (18S \text{ peak width} + 28S \text{ peak width}) - \text{threshold} = 0 \quad (3)$$

20 where a and b are the coefficient values determined by LDA.

[00325] It should be noted that in general, equation 3 can be rewritten as equation 4:

$$y = kx + b \quad (4)$$

in which a feature or a combination of features are used for the variables x and y . In this
 25 example, x is (18S area+28S area) and y is (18S width+28S width). The parameter k is the slope of the line which is determined by LDA. The parameter b is the intercept of the line which is the intersection point of the y axis and the line.

[00326] Multiple lines can be determined using multiple threshold values (each threshold value is associated with a line) and these lines can be used to create zones in

a plot of the two variables. For example, two lines can be determined for two different threshold values to divide the patients into three zones with two lines. This can be done such that the lines are parallel to one another as is shown in **FIG. 15B**. The slope of the lines is determined by the LDA and it depends on the actual data samples as well as the combination of features that is used for the different variables.

[00327] In order to determine the threshold values, many different threshold values are assessed in terms of one or more predictive or detective measures. For example, for each threshold value, LDA can be used to generate the line and then the samples in the resulting zones can be assessed to determine at least one of the sensitivity, specificity, negative predictive value, positive predictive value, false negative rate and false positive rate associated with the threshold value. The threshold value is then selected based on the one that has the best discriminative or predictive values. For example, the threshold values can be assessed on which threshold value has the best positive predictive value and the negative predictive value.

[00328] For each patient, using the two features that are selected as the x and y variable as well as the slope that is determined by LDA analysis, the equation 3 can be used to determine a particular intercept for the patient can be used as the LDA score for the patient.

[00329] For example, assume 0.1 and 0.2 are selected as the y-intercept to define the zone-partitions. The patients can then be divided into three zones as shown in **FIG. 15C**. The uppermost line is given by the equation $y = -1.6x + 0.2$ (or $8x + 5y = 1$) and the lowermost line is given by the equation $y = -1.6x + 0.1$ (or $8x + 5y = 0.5$). The x-axis and y-axis variables are (18S peak area+28S peak area) and (18S peak width+28S peak width), respectively.

[00330] The region underneath the lowermost line, above the x axis and to the right of the y axis defines zone 1 in which there are 29 patients, 6 of which are responders. The region between the uppermost and lowermost lines and to the right of the y axis defines zone 2 in which there are 22 patients, 2 of which are responders. The region above the uppermost line and to the right of the y-axis define zone 3 in which there are 34 patients, all of which are non-responders.

Example 15

[00331] In an alternative embodiment, rather than using LDA to determine RDA zones as was done in **FIGS. 15B** and **15C**, Quadratic Discriminant Analysis (QDA) can be used to determine partitions for the zones that can be used for prediction. With different thresholds, QDA gives different curves to divide the samples into various groups. For example, using two threshold values two groups, satisfying the following equation 5:

$$\begin{aligned}
 & c_1 \times (18S \text{ peak area} + 28S \text{ peak area})^2 + c_3 \times (18S \text{ peak width} + 28S \text{ peak width})^2 \\
 & + c_2 \times (18S \text{ peak area} + 28S \text{ peak area}) \times (18S \text{ peak width} + 28S \text{ peak width}) \\
 & + a_1 \times (18S \text{ peak area} + 28S \text{ peak area}) + b_1 \times (18S \text{ peak width} + 28S \text{ peak width}) - \\
 & \text{threshold} = 0
 \end{aligned} \tag{5}$$

where a first variable represents the feature 18S peak area+28S peak area and a second variable represents the feature 18S peak width+28S peak width and the parameters a_1 , b_1 , c_1 , c_2 and c_3 are determined by QDA.

[00332] Equation 4 can be used with two thresholds to generate two concentric curves that can divide the patients into three zones as shown in **FIG. 15D**. Once again, the circles 'o' represent responders, and the crosses 'x' represent non-responders. The appropriate threshold values can be determined as was explained previously using LDA, namely that for several threshold values the predictive measures of sensitivity, specificity, negative predictive value, positive predictive value, false negative rate and false positive rate can be calculated for each threshold. The desired threshold values can then be selected which result in the best predictive values.

[00333] For combinations of three features, the same analysis can be used to determine the RDA zones using either LDA or QDA.

[00334] It should be noted that in alternative embodiments, other techniques may be used to determine the RDA zones. For example, other machine learning methods may be used to do the prediction where there is a larger dataset such as, but not limited

- 88 -

to, Support Vector Machine (SVM), Neural Networks, and the k-nearest neighbors algorithm (kNN).

Example 16

5 RDI

[00335] The RNA Disruption index (RDI) is the output of an RDA assay described herein, for example a value corresponding to a ratio of at least some of the various combinations of features described herein or the result of LDA or quadratic discriminant analysis of at least two of the feature sets described herein. The cutoff point between
10 zone 2 and zone 3 may be chosen to maximize the PPV such that most of the responders from the MA22 dataset would be located in zone 3. For example, the RDI ranges used based on the study samples are about: 0.3 to 10 (zone 1), 10.1 to 35 (zone 2) and greater than 35.1 (zone 3). There may be other zones in other situations. For example, the fragmented RNA can be determined by the sum of the intermediate region
15 area and the low C region area. This index can be used to compare samples for the extent of RNA disruption, for example for research purposes. Clinical zones can be established which capture a range of RDI values wherein the clinical zones are defined by cutoff points (e.g. between zone 1 and 2 and between zone 2 and 3). For example the cutoff point between zone 1 and zone 2 can be set to correspond to a RDI value that
20 gives an NPV of 0.99.

Example 17

[00336] The data described in **Example 6** was reanalyzed using the peak identification method 400 described in **Example 10**.

25 **[00337]** Plots of the combination of features Intermediate Area/(28S+18S) and Low C Area/(28S+18S) were generated using the method 400 although method 300 may also be used.

- 89 -

[00338] The ratio of Low C Area/(28S+18S) was plotted against docetaxel dose. Little change was seen at 8hr, 24hr, or 48hr. However, after 72hr, a dramatic increase in the sample. Two replicate experiments had similar results.

[00339] The maximum of the ratio (Intermediate Area + Low C Area)/(28S+18S) was plotted against docetaxel dose and the results are shown in **FIG. 16B**. Each value represents a single sample. Two replicate experiments had similar results.

Example 18

[00340] Biopsy samples before and after a treatment dose were obtained from patients undergoing neo-adjuvant therapy with Trastuzumab. RNA was isolated and analyzed using the Agilent Bioanalyzer. Electropherograms were analyzed using the methods 100 and 400 and the maximal RDI score was graphed for each tumour sample at each time-point. The RDI values are increased in some patients after the first cycle of Trastuzumab (**FIG. 17A** and **FIG. 17B**).

[00341] Biopsy samples were obtained from patients undergoing neo-adjuvant therapy with Zometa. RNA was isolated and analyzed using the Agilent Bioanalyzer. Electropherograms were analyzed using the methods 100 and 400 and the maximal RDI score was graphed for each tumour sample at each time-point. The RDI values are increased in some patients after the first cycle of Zometa (**FIG. 17C** and **FIG. 17D**).

20

Example 19

[00342] **FIGS. 18A-18F** show several samples where the RIN algorithm gives a low RIN or "n/a" whereas the RDI score, determined as described herein, is better able to differentiate these samples. The electropherograms of **FIGS. 18A-18F** were taken from mid-therapy MA22 data. **FIGS. 18A-18F** also show that RIN values can be the same in electropherograms which demonstrate different patterns whereas the RDI score appears to be able to differentiate between these electropherograms. The RDI was calculated using the ratio or feature combination Intermediate Area/(28S+18S Areas).

25

[00343] FIG. 19 compares RIN to RDI in mid-therapy MA22 samples. When the RNA is intact, both the RIN and RDI scores are accurate. However, once the RNA becomes fragmented, for example due to drug exposure, the RDI is a better measure. The feature combination that was used for the RDI determination was Intermediate Area/(28S+18S).

5

Example 20

[00344] The need for biomarkers to accurately assess long term breast cancer chemotherapy efficacy is well established and is particularly needed for triple negative breast cancer (TNBC). To date, most in vitro diagnostic tests offered have been limited to hormone receptor positive patients. The RNA Disruption Assay (RDA), described herein, is a novel prognostic test for patients undergoing neoadjuvant chemotherapy that enables assessment of drug efficacy during treatment. RDA was developed in association with the CAN-NCIC-CTG MA-22 clinical trial for women with locally advanced breast cancer. Varying doses of docetaxel and epirubicin were given at two weekly (dose dense) or three weekly intervals. Tumours were biopsied in duplicate at three time points: pre-therapy, mid-therapy and post-therapy. Trial endpoints included clinical response, pathological complete response (pCR), disease-free survival and overall survival.

[00345] RDA is based on RNA electropherogram analysis generated to assess RNA quality as measured by the extent of ribosomal RNA (rRNA) disruption, mid-therapy. rRNA Disruption is quantified by a method described herein for example as described in FIGS. 5A and 5B or 12A and then stratified into an RNA Disruption Assay (RDA) score with 3 zones corresponding to the degree of RNA disruption. An RDA score in Zone 3 is associated with an increased likelihood of subsequent pathological complete response, an RDA score in Zone 2 is associated with an increased likelihood of at least partial response, and an RDA score in Zone 1 indicates that patients are very unlikely to receive long term chemotherapy benefit, NPV>0.98, for example.

[00346] When applied to the MA-22 study after the 3rd or 4th cycle of chemotherapy, 34 of 85 patients (40%) had tumours with RDA scores in zone 1.

[00347] Subsequently, it has been demonstrated in tumour cell lines that rRNA disruption leads directly to subsequent tumour cell death and that the extent of rRNA disruption is cell line type, dose, and time dependent.

5 Methods

[00348] Paraffin-embedded, formalin-fixed tumours from 82 MA-22 patients were subtyped by immunohistochemistry for estrogen receptor (ER), progesterone receptor (PR) and Her2/neu (Her2) expression status pre-therapy.

[00349] RDA, clinical response data and pCR occurrence data were obtained for each patient and matched to the tumour expression status for the above receptors. The feature combination that was used for the RDA determination was Intermediate Area/(28S+18S). The TNBC patients' tumours were then assessed for RNA disruption (RDA zones 1, 2 or 3) and compared to similar data for receptor positive patients. A similar comparison was made classifying patients by clinical response (e.g. no response/stable disease, partial response, or complete clinical response).

Results

[00350] Both Triple Negative Breast Cancer (TNBC) patients and receptor positive patients had tumours in each of the three RDA zones and in each of the clinical response categories. The pCR response rates were 3 of 21 (14% response rate) for TNBC patients, 0 of 43 (0% response rate) for ER+, Her2-ve patients, 1 of 6 (17% response rate) for ER+, Her2+ patients, and 4 of 12 (33% response rate) for Her2+ patients. Six of eight pCR responders were in RDA Zone 3 and two were in RDA Zone 2. Clinically, 6 of 8 pCR responders had a complete response, and two had partial responses. For the three TNBC clinical non responders who had stable disease, the RDA scores were in RDA Zone 3 for one patient, RDA Zone 2 for the second patient and RDA Zone 1 for the third. Seven of 18 non-responder TNBC patients were in RDA Zone 1. However, clinically, five of these patients had a partial response, one had a complete response and one had stable disease. Two of the three TNBC pCR responders showed an RDA score in RDA Zone 3 but clinically one had a partial response. For other

subtypes, 12 of 44 ER+, Her2-ve patients were in RDA Zone 1; 3 of 6 Er+ve, Her2+ve patients were in RDA Zone 1; 2 of 12 Her2+ve patients were in RDA zone 1. Using pCR and non pCR as definitive criteria, RDA was equally predictive with clinical response for prediction of complete response but was much superior for prediction of pCR non-
5 response.

Conclusions

[00351] In both TNBC and other subtypes of locally advanced breast cancer, RDA can identify pCR non-responders much better than clinical response. Currently, based
10 on this study, RDA can identify approximately 33% of TNBC patients as non-responders compared to approximately 28% for ER+ve Her2-ve patients. These results indicate that RDA, as a test for response guided therapy, can identify a subpopulation of non-responding TNBC patients who can be considered for alternate therapy.

15 Example 21

[00352] One major obstacle to the successful destruction of tumors using chemotherapy drugs is the presence of intrinsic or acquired resistance to anti-cancer agents. Assessed herein is whether chemotherapy-dependent alterations in tumor RNA quantity and integrity could also be demonstrated *in vitro* using the A2780 ovarian
20 cancer cell line. An equal number of cells were plated and treated with various docetaxel concentrations to determine the effect of the drug on cell division, cellular RNA content, and RNA integrity. At low docetaxel concentrations, RNA content increased per cell, likely due to increased rRNA production prior to a cell cycle arrest at mitosis. However, at higher docetaxel concentrations ($\geq 0.1 \mu\text{M}$), dose and time-dependent reductions in
25 cell number and cell RIN values were observed while the RDI based on the feature combination of Intermediate Area/(28S+18S) increased, which coincided with dramatic changes in the RNA banding pattern (i.e. RNA disruption). This included the formation of novel discrete bands distinct from the 28S and 18S rRNA bands. In contrast, treating docetaxel-resistant A2780DXL cells with docetaxel did not result in RNA disruption or
30 changes in cellular RNA content. These findings support the view that chemotherapy

dependent changes in tumor cell RNA content and integrity (as measured using the RNA disruption assay or RDA technique described herein) could effectively be used to monitor cellular response to chemotherapy agents and to differentiate between drug-sensitive and drug-resistant tumor cells *in vitro* and *in vivo*.

5 Example 22

[00353] Colon adenocarcinoma cells LS174T were incubated with Sunitinib, a receptor tyrosine kinase inhibitor. RNA was isolated from the cells and analyzed using an Agilent Bioanalyzer. RNA Disruption is evident at 24 hr at 54mM (see **FIG. 20A**).

[00354] Colon adenocarcinoma cells LS174T were incubated with Mafosfamide, a
10 DNA cross-linking agent. RNA was isolated from the cells and analyzed using an Agilent Bioanalyzer. RNA Disruption is evident at 48 hr at 48.2 mM (see **FIG. 20B**).

[00355] Colon adenocarcinoma cells SW480 were incubated with Epirubicin, a DNA synthesis inhibitor. RNA was isolated from the cells and analyzed using an Agilent Bioanalyzer. RNA Disruption is evident at 72 hr at 0.5 mM (see **FIG. 20C**).

15 **[00356]** It should be noted that either method 300 or method 400 can be used in determining the values for the combinations of features or the feature sets used in LDA and QDA analysis.

[00357] While the applicant's teachings described herein are in conjunction with various embodiments for illustrative purposes, it is not intended that the applicant's
20 teachings be limited to such embodiments. On the contrary, the applicant's teachings described and illustrated herein encompass various alternatives, modifications, and equivalents, without departing from the embodiments, the general scope of which is defined in the appended claims.

[00358] All publications, patents and patent applications are herein incorporated by
25 reference in their entirety to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference in its entirety. Specifically, the sequences associated with each accession number provided herein including, for example, accession numbers and/or biomarker sequences (e.g. protein and/or nucleic acid) provided in the Tables or elsewhere, are
30 incorporated by reference in its entirety.

CITATIONS FOR REFERENCES REFERRED TO IN THE SPECIFICATION

Brackstone et al Cancer Res 2012;72(24 Suppl):Abstract nr P1-14-13.

- 5 Cappallo-Obermann, H., W. Schulze, et al. (2011). "Highly purified spermatozoal RNA obtained by a novel method indicates an unusual 28S/18S rRNA ratio and suggests impaired ribosome assembly." Molecular Human Reproduction **17**(11): 669-678.

- 10 Fimognari C; Sestili P et al (2009) Protective effect of creatine against RNA damage
Mutation Research **670**:59-67

- 15 Hurley, J. (2011). "Meta-analysis of clinical studies of diagnostic tests: developments in how the receiver operating characteristic "works"." Arch Pathol Lab Med **135**(12): 1585-1590.

Li et al., Int. J. Mol. Med 2004 **14**(20): 257-264.

- 20 Mori, N., D. i. Mizuno, et al. (1978). "Increase in the ratio of 18S to 28S RNA in the cytoplasm of mouse tissues during aging." Mechanisms of Ageing and Development: 285-297.

Mueller, O. L., S.;Schroeder, A (2004) RNA Integrity Number (RIN) – Standardization of RNA Quality Control.

- 25 Parissenti, A. M., J.-A. W. Chapman, et al. (2010). "Association of low tumor RNA integrity with response to chemotherapy in breast cancer patients." Breast Cancer Res Treat **119**: 347-356.

- 30 Schroeder, A., O. Mueller, et al. (2006). "The RIN: an RNA integrity number for assigning integrity values to RNA measurements." BMC Mol Biol **7**: 3.

Skrypina, N. A., A. V. Timofeeva, et al. (2003). "Total RNA suitable for molecular biology analysis." Journal of Biotechnology **105**: 1-9.

- 35 Vespucci, U. G. (2005). Agilent 2100 Bioanalyzer, Agilent Technologies, Inc. 2100 Expert User's Guide. Agilent Technologies Hewlett-Packard Str. 8
76337 Waldbronn Germany, Agilent Technologies, Inc. . Agilent2100 G2946-90004_Vespucci_UG_eBook_(NoSecPack)[2].

CLAIMS:

1. A method for performing an RNA Disruption Assay (RDA) for cellular RNA optionally in response to a cytotoxic treatment optionally a treatment for a proliferative disorder optionally a cancer treatment, optionally chemotherapy, adjuvant or radiation
5 treatment, wherein the method comprises:
obtaining at least one electropherogram dataset corresponding to a unique biological sample comprising cellular RNA optionally at a time point before, during or after the treatment;
determining values for features from the at least one electropherogram
10 dataset by using two identifying ranges to accommodate possible shifting of 18S and 28S peaks, detecting the 18S and 28S peaks, and calculating the features derived at least in part based on the located 18S and 28S peaks; and
optionally determining an RDA score based on a combination of the values of the features.
- 15 2. The method of claim 1, wherein the method comprises:
obtaining two electropherogram datasets corresponding to a test sample and a control sample, the test sample being a unique biological sample comprising cellular RNA optionally at a time point before, during or after the treatment requiring analysis and the control sample being another unique biological sample or a positive control sample
20 of intact RNA or approximately intact RNA;
placing the electropherogram of the test sample into an adjustment group and placing the electropherogram of the control sample into a standard group;
identifying normal characteristics representative of a normal sample from the standard group;
25 determining locations of peaks for the electropherogram in the adjustment group and adjusting the location of the peaks based on the normal characteristics; and
determining values for features from the adjusted electropherogram and optionally the electropherogram from the standard group when the control sample is another unique biological sample requiring analysis.
- 30 3. The method of claim 1, wherein the method comprises:

obtaining a plurality of electropherogram datasets corresponding to unique biological samples comprising cellular RNA optionally at a time point before, during or after the treatment;

5 separating the plurality of electropherograms into an adjustment group and a standard group;

identifying normal characteristics representative of a normal sample from the plurality of electropherograms in the standard group;

determining locations of peaks for the electropherograms in the adjustment group and adjusting the location of the peaks based on the normal characteristics; and

10 determining values for features from the adjusted electropherograms and the electropherograms from the standard group.

4. The method of claims 1, 2 or 3, wherein the method comprises defining default ranges as the two identifying ranges to search for the 18S and 28S peaks and adjusting the default ranges if a marker region is not at its expected location.

15 5. The method of claim 4, wherein the default ranges comprise about 39.5 seconds to about 44.95 seconds for the 18S peak and about 45.05 seconds to 53.5 seconds for the 28S peak.

6. The method of any one of claims 2 to 5, wherein the default ranges are adjusted according to $rngshft = marker\ time - 22.5$ if a start time of the marker region is not at
20 about 22.5 seconds.

7. The method of claim 6, wherein the default range for the 18S peak is shifted according to $[39.5 + rngshft \times 1.5, 44.95 + rngshft \times 1.5]$ when the marker region is not at its expected location.

8. The method of claims 6 or 7, wherein the default range for the 28S peak is
25 shifted according to $[45.05 + rngshft \times 2.5, 53.5 + rngshft \times 2.5]$.

9. The method of any one of claims 2 to 8, wherein peaks are located by first locating possible peak candidates in the two identifying ranges, calculating height

differences of each possible peak candidate, and defining the 18S and 28S peaks as the two largest peak candidates.

10. The method of claim 9, wherein a height difference of a given peak candidate is defined as a difference of a highest peak value and an average of lowest peak values for the given peak candidate where the low peak values that are averaged are a starting point and an ending point associated with the highest peak value.

11. The method of any one of claims 2 to 10, wherein, based on the 18S and 28S peak locations, values for features can be calculated comprising at least one of 18S peak width, 18S peak height, 18S peak area, 28S peak width, 28S peak height, 18S peak width, intermediate area, and low C banding region area.

12. The method of any one of claims 3 to 11, wherein the separating act comprises determining a standard sample for the plurality of electropherogram datasets where the standard sample has the smallest standard score defined by a ratio of a sum of intermediate area, lower B banding region area and lower C banding region area to a sum of 18S peak area and 28S peak area.

13. The method of claim 12, wherein the separating act further comprises assigning a given electropherogram dataset to the adjustment group if retention times for at least one of the 18S and 28S peaks of the given sample differ by more than a certain time threshold to retention times of corresponding at least one of the 18S and 28S peaks of the standard sample.

14. The method of claim any one of claims 3 to 13, wherein the identifying act comprises defining standard retention times as the normal characteristics, the standard retention times being defined by a statistical measure of corresponding retention times for electropherograms in the standard group.

15. The method of claim 14, wherein the statistical measure comprises one of a median and a mean.

16. The method of any one of claims 2 to 15, wherein the act of determining locations of peaks for the electropherograms in the adjustment group comprises

determining retention times for 18S and 28S peaks of the electropherograms in the adjustment group by locating peaks around corresponding standard retention times.

17. The method of any one of claims 3 to 16, wherein the plurality of electropherogram datasets are derived from unique biological samples on an RNA chip.

5 18. The method of any one of claims 1 to 17, wherein the RDA score comprises at least two RDA zones.

19. The method of any one of claims 1 to 18, wherein the RDA score comprises an RDI score.

10 20. The method of any one of claims 2 to 19, wherein the features are used to define a ratio of low C banding region area to a sum of the 18S peak area and the 28S peak area, where the low C banding region is a portion of an intermediate region below the 18S peak region.

15 21. The method of any one of claims 2 to 19, wherein the features are used to define a ratio of the sum of the low C banding region area and intermediate area to the sum of the 28S peak area and the 18S peak area.

22. The method of any one of claims 2 to 19, wherein the features are used to define a ratio of the sum of the low C banding region area and the intermediate area to the sum of the 28S peak area and the 18S peak area

20 23. The method of any one of claims 2 to 19, wherein the features are used to define a maximum ratio determined from at least two electropherogram datasets, wherein the ratio is defined as an intermediate area divided by the sum of the 18S and 28S peak areas.

25 24. The method of any one of claims 2 to 19, wherein the features are used to define a ratio of low C banding region area to a sum of the 18S peak area and the 28S peak area.

25. The method of any one of claims 2 to 19, wherein the features are used to define a ratio of intermediate area to a sum of the 28S peak area and the 18S peak area.
26. The method of any one of claims 2 to 19, wherein three dimensional feature sets
5 are used to determine the RDA score, the three dimensional feature sets each comprising at least one feature and each feature set being assigned to an orthogonal variable, the RDA score being compared to an RDA plane for analysis.
27. The method of claim 26, wherein the first variable comprises a sum of 18S peak area and 28S peak area, the second variable comprises a sum of 18S peak height and
10 28S peak height and the third variable comprises low C banding area.
28. The method of claim 26, wherein the first variable comprises a sum of 18S peak area and 28S peak area, the second variable comprises a sum of 18S peak height and 28S peak height and the third variable comprises a sum of 18S peak width and 28S peak width.
- 15 29. The method of claim 26, wherein the first variable comprises a sum of 18S peak area and 28S peak area, the second variable comprises a sum of 18S peak height and 28S peak height and the third variable comprises intermediate area.
30. The method of claim 26, wherein the first variable comprises a sum of 18S peak area and 28S peak area, the second variable comprises a sum of 18S peak width and
20 28S peak width and the third variable comprises low C band area.
31. The method of any one of claims 2 to 19, wherein two dimensional feature sets are used to determine the RDA score, the two dimensional feature sets comprising at least one feature in each feature set and each feature set being assigned to an orthogonal variable, the RDA score being compared to an RDA line or RDA zone for
25 analysis.
32. The method of claim 31, wherein the first variable comprises 18S peak height and the second variable comprises 28S peak height.

33. The method of claim 31, wherein the first variable comprises a sum of the 28S peak area and the 18S peak area and the second variable comprises low C banding region area.
34. The method of claim 31, wherein the first variable comprises a sum of the 28S
5 peak area and 18S peak area and the second variable comprises low B banding region area.
35. The method of claim 31, wherein the first variable comprises the sum of 18S peak area and 28S peak area and the second variable comprises a sum of 18S and 28S peak widths.
- 10 36. The method of claim 31, wherein the first variable comprises a sum of the 28S and 18S peak heights and the second variable comprises a sum of the 28S peak area and the 18S peak area.
37. The method of claim 31, wherein the variable comprises a sum of 18S peak area and 28S peak area and the second variable comprises intermediate area.
- 15 38. The method of claim 31, wherein the first variable comprises a sum of 18S peak area and 28S peak area and the second variable comprises an RNA concentration of a corresponding sample.
39. The method of any one of claim 31, wherein the first variable comprises a sum of the 18S and 28S peak heights and the second variable comprises an RNA concentration
20 of a corresponding sample.
40. The method of claim 31, wherein the first variable comprises a sum of 18S and 28S peak heights and the second variable comprises intermediate area.
41. The method of claim 31, wherein the first variable comprises a sum of the 18S and 28S peak heights and the second variable comprises low C banding region area.

42. The method of claim 31, wherein the first variable comprises a sum of 18S and 28S peak widths and the second variable comprises a sum of 18S peak area and the 28S peak area.
43. The method of any one of claims 2 to 19, wherein a single feature is used to
5 determine the RDA score, the single feature comprising the 18S peak area and the single feature being compared to an RDA threshold or control for analysis.
44. The method of claim 43, wherein the single feature comprises the 28S peak area.
45. The method of claim 43, wherein the single feature comprises 18S peak height.
- 10 46. The method of claim 43, wherein the single feature comprises 28S peak height.
47. The method of claim 43, wherein the single feature comprises intermediate area between 18S and 28S regions.
48. The method of claim 43, wherein the single feature comprises low C banding region area.
- 15 49. The method of claim 43, wherein the single feature comprises low B banding region area.
50. The method of claim 1, wherein allowing for possible shifting of the 18S and 28S peaks comprises defining at least two shifted regions of the at least one electropherogram dataset and determining values for features from the at least two
20 shifted regions.
51. The method of claim 50, wherein the shifted regions comprise at least one of a shifted 18S region, a shifted 28S region, an intermediate banding region between an 18S peak in the shifted 18S region and a 28S peak in the shifted 28S region and a low banding region below the 18S peak excluding a marker region.
- 25 52. The method of claim 51, wherein the combination of features comprises a ratio of 28S peak area in the shifted 28S region to 18S peak area in the shifted 18S region.

53. The method of claim 51, wherein the combination of features comprises a ratio of 28S peak area in the shifted 28S region to total area, the total area comprising 18S peak area in the shifted 18S region, the 28S peak area, intermediate banding region area and low banding region area.
- 5 54. The method of claim 51, wherein the combination of features comprises a ratio of 18S peak area in the shifted 18S region to total area, the total area comprising the 18S peak area, 28S peak area in the shifted 28S region, intermediate banding region area and low banding region area.
- 10 55. The method of claim 51, wherein the combination of features comprises a ratio of intermediate banding region area to total area, the total area comprising 18S peak area in the shifted 18S region, 28S peak area in the shifted 28S region, the intermediate banding region area and low banding region area.
- 15 56. The method of claim 51, wherein the combination of features comprises a ratio of low banding region area to total area, the total area comprising 18S peak area in the shifted 18S region, 28S peak area in the shifted 28S region, intermediate banding region area and the low banding region area.
- 20 57. The method of claim 51, wherein the combination of features comprises a ratio of intermediate banding region area to the addition of 18S peak area in the shifted 18S region and 28S peak area in the shifted 28S region.
- 25 58. The method of claim 51, wherein the combination of features comprises a ratio of low banding region area to the addition of 18S peak area in the shifted 18S region and 28S peak area in the shifted 28S region.
59. The method of claim 51, wherein the combination of features comprises a ratio of intermediate banding region area to 18S peak area in the shifted 18S region.
60. The method of claim 51, wherein the combination of features comprises a ratio of low banding region area to 18S peak area in the shifted 18S region.

61. The method of claim 51, wherein the combination of features comprises a ratio of intermediate banding region area to 28S peak area in the shifted 28S region.
62. The method of claim 51, wherein the combination of features comprises a ratio of low banding region area to 28S peak area in the shifted 28S region.
- 5 63. The method of claim 51, wherein the combination of features comprises a ratio of the addition of low banding region area and intermediate banding region area to the addition of 18S peak area in the shifted 18S region and 28S peak area in the shifted 28S region.
- 10 64. The method of claim 51, wherein the combination of features comprises a maximum ratio determined from at least two electropherogram datasets, the ratio being defined as intermediate banding region area divided by a sum of 18S peak area in the shifted 18S region and 28S peak area in the shifted 28S region.
- 15 65. The method of claim 51, wherein the combination of features comprises a minimum ratio determined from at least two electropherogram datasets, the ratio being defined as intermediate banding region area divided by a sum of 18S peak area in the shifted 18S region and 28S peak area in the shifted 28S region.
- 20 66. The method of claim 51, wherein the combination of features comprises a maximum ratio determined from at least two electropherogram datasets, the ratio being defined as intermediate banding region area divided by total area, the total area being a sum of 18S peak area in the shifted 18S region, 28S peak area in the shifted 28S region, intermediate banding region area and low banding region area.
- 25 67. The method of claim 51, wherein the combination of features comprises a minimum ratio determined from at least two electropherogram datasets, the ratio being defined as low banding region area divided by total area, the total area being a sum of 18S peak area in the shifted 18S region, 28S peak area in the shifted 28S region, intermediate banding region area and low banding region area.
68. The method of claim 51, wherein the combination of features comprises a ratio of low banding region area divided by total area and a maximum second ratio determined

from at least two electropherogram datasets, the second ratio being defined as intermediate banding area divided by total area, the total area being a sum of 18S peak area in the shifted 18S region, 28S peak area in the shifted 28S region, intermediate banding region area and the low banding region area.

- 5 69. The method of claim 51, wherein the combination of features comprises a maximum ratio determined from at least two electropherogram datasets, the ratio being defined as 28S peak area divided by total area, the total area being a sum of 18S peak area in the shifted 18S region, 28S peak area in the shifted 28S region, intermediate banding region area and low banding region area.
- 10 70. The method of claim 51, wherein the combination of features comprises a minimum ratio determined from at least two electropherogram datasets, the ratio being defined as 28S peak area divided by total area, the total area being a sum of 18S peak area in the shifted 18S region, 28S peak area in the shifted 28S region, intermediate banding region area and low banding region area.
- 15 71. The method of claim 51, wherein the combination of features comprises a maximum ratio determined from at least two electropherogram datasets, the ratio being defined as 18S peak area divided by total area, the total area being a sum of 18S peak area in the shifted 18S region, 28S peak area in the shifted 28S region, intermediate banding region area and low banding region area.
- 20 72. The method of claim 51, wherein the combination of features comprises a minimum ratio determined from at least two electropherogram datasets, the ratio being defined as 18S peak area divided by total area, the total area being a sum of 18S peak area in the shifted 18S region, 28S peak area in the shifted 28S region, intermediate banding region area and low banding region area.
- 25 73. The method of claim 53, wherein the combination of features comprises a combination of an 18S peak width and a 28S peak width.
74. The method of claim 52, wherein the shifted regions comprise a shifted 18S region in the range of 40.9 to 50.4 units and a shifted 28S region in the range of 58.1 to

69.5 units when a time axis of the at least one electropherogram dataset is converted to a scale of 0 to 100 units.

75. The method of claim 74, wherein if an 18S peak area in the shifted 18S region divided by total area multiplied by 100 is less than about 8 times a 28S peak area in the shifted 28S region divided by total area multiplied by 100, the shifted 18S region is defined to be the range of 42.8 to 51.4 units, the total area comprising the 18S peak area, the 28S peak area, intermediate banding region area and low banding region area.

76. The method of claim 74 or claim 75, wherein an 18S peak width is between 0.6 and 7.6 units.

77. The method of any one of claims 74 to 76, wherein if a 28S peak area in the shifted 28S region divided by total area multiplied by 100 is less than about 8 times an 18S peak area in the shifted 18S region divided by total area multiplied by 100, the shifted 28S region is defined to be the range of 52.3 to 69.5 units, the total area comprising the 18S peak area, the 28S peak area, intermediate banding region area and low banding region area.

78. The method of any one of claims 74 to 77, wherein a 28S peak width is between 0.8 and 10.5 units.

79. The method of claim 78, wherein if the 18S peak area is less than a threshold, the 18S peak is defined to comprise more than one maximum value and wherein if the 28S peak area is less than a threshold, the 28S peak is defined to comprise more than one maximum value.

80. The method of claim 79, wherein the threshold is 0.2 FU-seconds.

81. The method of claim 50, wherein the method further comprises defining boundaries for a shifted 18S region and a shifted 28S region of the at least one electropherogram dataset, determining peak area of an 18S peak in the shifted 18S region and a 28S peak in the shifted 28S region, determining if the 18S peak and the 28S peak should have more than one maximum value depending on the 18S peak area and the 28S peak area, and determining if boundaries of the shifted 18S region and the

shifted 28S region require adjustment depending on the areas of the 18S peak and 28S peak with respect to each other.

82. The method of claim 81, wherein the method further comprises adjusting widths of the 18S peak and the 28S peak to be within a certain width range, and adjusting a
5 location of at least one of the 18S peak and the 28S peak if a distance between the 18S peak and the 28S peak is greater than a distance threshold.

83. The method of any one of claims 74 to 82, wherein an ending position of an 18S peak and a starting position of a 28S peak are less than 17.1 units apart.

84. The method of any one of claims 1 to 83, wherein a peak area for a given peak is
10 defined by an area between two local minimum values on either side of the given peak.

85. The method of any one of claims 1 to 83, wherein the step of obtaining at least one electropherogram dataset comprises: separating cellular RNA optionally cancerous tumour cell RNA obtained from a subject by electrophoresis; detecting the quantity of separated cellular RNA optionally cancerous tumour cell RNA; plotting the quantity of
15 separated cellular RNA according to obtain an electropherogram and obtaining the electropherogram dataset from the electropherogram.

86. The method of claim 85, wherein the gel electrophoresis comprises microcapillary gel electrophoresis.

87. The method of claim 85, wherein the electrophoresis comprises using an RNA
20 chip comprising microchannels for electrophoretically separating the RNA of a given sample.

88. The method of any one of claims 1 to 87, wherein the method further comprises calculating an RDA score based on a combination of the values of the features, comparing the RDA score to a threshold optionally to assess effectiveness of the
25 treatment.

89. The method of any one of claim 88, wherein the method further comprises comparing the RDA score to multiple thresholds, each threshold arranged to

corresponding to an RDA zone boundary to determine effectiveness of the treatment, optionally wherein one of the thresholds corresponds to a selected negative predictive value (NPV) and defines a boundary between RDA zone 2 and RDA zone 3 and another of the threshold corresponds to a selected positive predictive value (PPV) and defines the boundary between a RDA zone 1 and RDA zone 2.

90. The method of claim 89, wherein the NPV is at least 0.98 and/or the PPV is at least 0.18, 0.19 or 0.20.

91. The method of any one of claims 1 to 89, wherein the cellular RNA is cancerous tumour cell RNA.

10 92. An assay comprising:

quantifying the amount of RNA disruption in a biological sample comprising cellular RNA at a time point before, during or after cytotoxic treatment, optionally a cancer treatment, optionally a chemotherapeutic, adjuvant and/or radiation treatment, using an RNA disruption assay (RDA) performed according to the method of any one of claims 1 to 90; and

determining if the amount of RNA disruption of the cellular RNA is increased or not increased.

93. The assay of claim 92, wherein the determining step comprises comparing the quantified amount of RNA disruption in the biological sample with a threshold or reference value.

94. The assay of claim 92 or 93, wherein the cellular RNA is cancerous tumour cell RNA.

95. The assay of claim 94, further comprising:
identifying the tumour as responsive to the chemotherapeutic treatment if the amount of RNA disruption is increased relative to the threshold or reference value.

96. The assay of claim 95, wherein the biological sample corresponds to cancerous tumour cells treated in vitro or in vivo.

97. An in vitro assay for determining sensitivity of cells to a cytotoxic treatment optionally a chemotherapeutic, adjuvant and/or radiation treatment, the assay comprising:

- 5 quantifying the amount of RNA disruption in a biological sample comprising cell RNA at a time point before, during or after treatment using an RNA Disruption Assay (RDA) performed according to the method of any one of claims 1 to 90; and
- 10 comparing the quantified amount of RNA disruption with a threshold or reference value or a threshold or reference two or three dimensional region, and if the amount of RNA disruption is increased relative to the reference value, identifying the cell as sensitive, and if not increased identifying the cells as resistant to the treatment.

98. An assay for determining/predicting if a subject is responding to a cytotoxic therapy, optionally a cancer therapy, optionally a chemotherapeutic drug and/or radiation
- 15 treatment comprising: assaying a biological sample obtained from the subject before, during and/or after the subject has received a treatment for the quantity of RNA disruption using an RNA Disruption Assay (RDA) performed according to the method of any one of claims 1 to 90, wherein the subject is identified as responding to the treatment if the quantity of RNA disruption is increased relative to the threshold or
- 20 reference value.

99. The assay of claim 98, wherein the response is indicative of treatment outcome, and an increased quantity of non-autolytic RNA disruption relative to a threshold or reference value is indicative of a positive treatment outcome and a lack of non-autolytic RNA disruption relative to a threshold or reference value is indicative of a negative
- 25 treatment response.

100. The assay of claim 98, wherein the positive treatment outcome is pathologic complete response following treatment, partial response, reduced risk of disease progression, stable disease or increased overall survival.

101. The assay of any one of claims 98 to 100, wherein the biological sample is obtained after the subject has received a first cycle of a chemotherapeutic, adjuvant and/or radiation treatment.
102. The assay of any one of claims 98 to 100, wherein the biological sample is
5 obtained from the subject after the subject has received 2, 3, 4 or more cycles of the chemotherapeutic, adjuvant and/or radiation treatment.
103. The assay of any one of claims 98 to 100, wherein the biological sample is obtained from the subject after completion of the chemotherapeutic, adjuvant and/or radiation treatment regimen.
- 10 104. The assay of any one of claims 91 to 102, where RNA disruption score indicative of responsiveness and/or positive treatment outcome correlates to a decrease in RNA integrity of at least by 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80% or at least 90%.
105. The assay of any one of claims 91 to 102, wherein the reference value is a
15 pretreatment value.
106. A method of treating a subject in need of cancer treatment, the method comprising:
administering a cytotoxic treatment, optionally a chemotherapeutic drug, adjuvant and/or a radiation treatment;
20 determining if the subject is responding to the cancer treatment according to the method of any one of claims 98 to 105; and
continuing administration of the cytotoxic treatment if the subject is responsive and/or discontinuing and/or altering the treatment if the subject is not responsive.
- 25 107. The method of claim 106, wherein the cancer treatment is a chemotherapeutic treatment selected from microtubule stabilizing agents optionally Docetaxel and paclitaxel, DNA synthesis inhibitors optionally Epirubicin, inhibitors of Her2 Receptor optionally Trastuzumab, DNA cross-linking agents optionally Mafosfamide, carboplatin

and cisplatin, VEGFA inhibitors optionally Bevacizumab, Receptor Tyrosine Kinase inhibitors optionally Sunitinib and Toceranib, Bisphosphonates optionally Zoledronic acid, Thymidylate synthase inhibitors optionally 5-fluorouracil, taxanes, anthracyclines, and vinca alkaloids optionally vinblastine, and nucleoside analogs optionally 5-FU.

5 108. The method of claim 106, wherein the cytotoxic treatment is selected from docetaxel, paclitaxel, epirubicin, doxorubicin, vinblastine, cisplatin and combinations thereof.

109. The method of any one of claims 1 to 90 and 106 or the assay of any one of claims 91 to 104, wherein the cancer is selected from breast cancer, prostate cancer,
10 lung cancer, sarcomas, colon cancer, leukemia, lymphoma, multiple myeloma or ovarian cancer.

110. The method or assay of claim 109, wherein the cancer is breast cancer.

111. The method or assay of claim 109 or 110, wherein the breast cancer is HER2+ positive breast cancer.

15 112. The method or assay of claim 109 or 111 wherein the cancer treatment comprises chemotherapy treatment selected from taxanes, anthracyclines optionally selected from one or more of paclitaxel, docetaxel, docetaxel and epirubicin and combinations thereof.

20 113. The method or assay of claim 112, wherein the taxane is selected from paclitaxel and docetaxel and/or the anthracycline is selected from doxorubicin and epirubicin.

114. The method or assay of claim 109, wherein the cancer is ovarian cancer.

115. The method or assay of claim 114, wherein the cancer therapy is a chemotherapeutic treatment selected from taxanes, anthracyclines, and vinca alkaloids optionally vinblastine, alkylating agents optionally cisplatin and nucleoside analogs
25 optionally 5-FU and combinations thereof.

116. The method or assay of claim 115, wherein the taxane is selected from paclitaxel and/or docetaxel and combinations thereof and/or the anthracycline is selected from doxorubicin and epirubicin and combinations thereof.

117. A non-transitory computer readable medium comprising a plurality of instructions
5 executable on a microprocessor of an apparatus for adapting the apparatus to implement a method for performing an RNA Disruption Assay (RDA) for cellular RNA optionally a treatment for a proliferative disorder optionally a cancer treatment, optionally chemotherapy, adjuvant or radiation treatment, wherein the method comprises:

- accessing at least one electropherogram dataset corresponding to a unique
10 biological sample comprising cellular RNA optionally at a time point before, during or after the treatment;
- determining values for features from the at least one electropherogram dataset by using two identifying ranges to accommodate possible shifting of 18S and 28S peaks, detecting the 18S and 28S peaks, and calculating the features derived at
15 least in part based on the located 18S and 28S peaks; and
- optionally determining an RDA score based on a combination of the values of the features.

118. The non-transitory computer readable medium of claim 116, wherein the method is further defined according to any one of claims 2 to 90.

20 119. An apparatus for performing an RNA Disruption Assay (RDA) for cellular RNA optionally a treatment for a proliferative disorder optionally a cancer treatment, optionally chemotherapy adjuvant or radiation treatment, wherein the apparatus comprises:

- a data input to receive at least one electropherogram dataset corresponding to a unique biological sample comprising cellular RNA optionally at a time point before,
25 during or after the treatment;
- a processing unit coupled to the data input, the processing unit being configured to determine values for features from the at least one electropherogram dataset by using two identifying ranges to accommodate possible shifting of 18S and 28S peaks, detecting the 18S and 28S peaks, and calculating the features derived at
30 least in part based on the located 18S and 28S peaks; and

a data output coupled to the processing unit to convey an indication of the RDA score.

120. The apparatus of claim 118, wherein the processing unit is configured to perform the method of any one of claims 2 to 90.

5 121. The apparatus of claim 119 or claim 120, wherein the apparatus further comprises an electrophoresis device coupled to the data input, the electrophoresis device being configured to generate the at least one electropherogram dataset.

122. An apparatus for performing an RNA Disruption Assay (RDA) for cellular RNA optionally a treatment for a proliferative disorder optionally a cancer treatment, optionally
10 chemotherapy, adjuvant or radiation treatment, wherein the apparatus comprises:

means for obtaining at least one electropherogram dataset corresponding to a unique biological sample comprising cellular RNA optionally at a time point before, during or after the treatment;

15 means for determining values for features the at least one electropherogram dataset by using two identifying ranges to accommodate possible shifting of 18S and 28S peaks, detecting the 18S and 28S peaks, and calculating the features derived at least in part based on the located 18S and 28S peaks; and

means for determining an RDA score based on a combination of the values of the features.

20 123. The apparatus of claim 122, wherein the means for determining values for features from at least two shifted regions of the electropherogram data is configured to perform the method defined in any one of claims 2 to 17 and 73 to 84.

124. The apparatus of claim 122 or claim 123, wherein the means for determining the RDA score is configured to perform the method defined in any one of claims 18 to 73, 88
25 and 89.

125. A method for performing an RNA Disruption Assay (RDA) for cellular RNA, wherein the method comprises:

obtaining at least one electropherogram dataset corresponding to a unique biological sample comprising the cellular RNA at a time point;

defining an 18S peak and a 28S peak from the at least one electropherogram dataset;

5 determining at least one parameter value for both the 18S peak and the 28S peak;

redefining at least one of the 18S peak and the 28S peak when required according to one or more rules applied to the at least one parameter value;

determining an 18S peak area and a 28S peak area; and

10 determining an RDA score based on at least one of the 18S peak area and the 28S peak area.

126. The method of claim 125, wherein the at least one parameter value comprises at least one of a peak area, a peak width and a peak location.

127. The method of claim 125, wherein the one or more rules comprises at least one
15 of determining if at least one of the 18S peak and the 28S peak have more than one maximum value based on peak area, determining if boundaries where the 18S peak and the 28S peak are located require adjustment depending on comparing the areas of the 28S peak and 18S peak to one another, adjusting widths of the 18S peak and the 28S peak to be in a certain width range, and adjusting a location of at least one of the 18S
20 peak and the 28S peak if a distance between the 18S peak and 28S peak is greater than a distance threshold.

128. A method for performing an RNA Disruption Assay (RDA) for cellular RNA, wherein the method comprises:

25 obtaining at least one electropherogram dataset corresponding to a unique biological sample comprising the cellular RNA at a time point;

defining an 18S peak and a 28S peak from the at least one electropherogram dataset;

determining at least one parameter value for both the 18S peak and the 28S peak; and

redefining at least one of the 18S peak and the 28S peak when required according to one or more rules applied to the at least one parameter value.

129. The method of claim 128, wherein the at least one parameter value comprises at least one of a peak area, a peak width and a peak location.

5 130. The method of claim 128, wherein the one or more rules comprises at least one of determining if at least one of the 18S peak and the 28S peak have more than one maximum value based on peak area, determining if boundaries where the 18S peak and the 28S peak are located require adjustment depending on comparing areas of the 28S peak and 18S peak to one another, adjusting widths of the 18S peak and the 28S peak
10 to be in a certain width range, and adjusting a location of at least one of the 18S peak and the 28S peak if a distance between the 18S peak and 28S peak is greater than a distance threshold.

131. The method of claim 128, wherein the RDA is performed optionally in response to a cytotoxic treatment optionally a treatment for a proliferative disorder optionally a
15 cancer treatment, optionally chemotherapy, adjuvant or radiation treatment , and the method comprises:

obtaining the at least one electropherogram dataset corresponding to the unique biological sample comprising the cellular RNA during or after the treatment;
determining values for features from at least two shifted regions of the at
20 least one electropherogram dataset, the shifting being due to the treatment; and
determining the RDA score based on a combination of the values of the features.

132. The method of claim 131, wherein the shifted regions comprise at least one of a shifted 18S region, a shifted 28S region, an intermediate banding region between the
25 18S peak in the shifted 18S region and the 28S peak in the shifted 28S region and a low banding region below the 18S peak excluding a marker region.

133. The method of claim 132, wherein the method further comprises performing the RDA according to any one of claims 52 to 91.

134. A non-transitory computer readable medium comprising a plurality of instructions executable on a microprocessor of an apparatus for adapting the apparatus to implement a method for performing an RNA Disruption Assay (RDA) for cellular RNA, wherein the method comprises:

- 5 obtaining at least one electropherogram dataset corresponding to a unique biological sample comprising the cellular RNA at a time point;
 - defining an 18S peak and a 28S peak from the at least one electropherogram dataset;
 - determining at least one parameter value for both the 18S peak and the
- 10 28S peak; and
 - redefining at least one of the 18S peak and the 28S peak when required according to one or more rules applied to the at least one parameter value.

135. The non-transitory computer readable medium of claim 134, wherein the method is further defined according to any one of claims 129 to 133.

- 15 136. An apparatus for performing an RNA Disruption Assay (RDA) for cellular RNA, wherein the apparatus comprises:
 - a data input to receive at least one electropherogram dataset corresponding to a unique biological sample comprising cellular RNA at a time point;
 - a processing unit coupled to the data input, the processing unit being
- 20 configured to define an 18S peak and a 28S peak from the at least one electropherogram dataset; determine at least one parameter value for both the 18S peak and the 28S peak; and redefine at least one of the 18S peak and the 28S peak when required according to one or more rules applied to the at least one parameter value; and
 - a data output coupled to the processing unit to convey an indication of the
- 25 RDA.

137. The apparatus of claim 136, wherein the processing unit is configured to perform the method of any one of claims 129 to 133.

- 138. The apparatus of claim 135 or claim 136, wherein the apparatus further comprises an electrophoresis device coupled to the data input, the electrophoresis
- 30 device being configured to generate the at least one electropherogram dataset.